FULL SPEED AHEAD

HOW TO MAKE VARIABLE DATA PDF FILES THAT WON'T SLOW YOUR DIGITAL PRESS

Edited by Global Graphics

SECOND EDITION

GLOBAL GRAPHICS®

With thanks to our sponsors...

Indigo WhatTheyThink DIGIMARC DELPHAX SOLUTIONS racami KODAK HYBRID SOFTWARE PageWide Industrial

WITH THANKS TO:

The contents of this document are provided "AS IS" and without warranty. This document could contain technical inaccuracies, typographical errors and out-of-date information. The information contained herein may be updated or changed without notice at any time. Use of the information is therefore at your own risk. Global Graphics shall not be liable for technical or editorial errors or omissions contained herein.

For educational purposes only

This document may be used and reproduced by educators and students solely for purposes directly related to academic work.

First published by Global Graphics PLC in March 2020.
Second edition published January 2021.

You may not edit or modify this document without the prior written approval of Global Graphics.

Global Graphics, Harlequin, the Harlequin logo, the Harlequin RIP, are trademarks of Global Graphics Software Limited which may be registered in certain jurisdictions. Global Graphics is a trademark of Global Graphics PLC which may be registered in certain jurisdictions. PostScript, Adobe and Photoshop are trademarks of Adobe Systems Incorporated which may be registered in certain jurisdictions. All other brand and product names are the registered trademarks or trademarks of their respective owners.

Companies in the Global Graphics group are: Global Graphics Software; Meteor Inkjet; Xitron; and, HYBRID Software.

CONTENTS

© Global Graphics 2021. All rights reserved.

FOREWORD

by Bob Leahey and Pat McGrew

Full Speed Ahead should be read by everyone in the digital print supply chain. Designers will learn why the file is structured as it is, how layers are used, and how the use of variable data will affect print on a digital press. Customer service team members will learn how to spot files that could pose challenges in the production workflow. Prepress teams will have the ammunition they need to recommend best practices to clients. Operators will learn why some files behave differently than others, even when they appear quite similar. And creators of software tools will learn why the creation of the most efficient PDF possible should be their goal. This book also provides the information needed to create that elusive electronic nugget, the optimally efficient file.

In his own introduction to this book, author Martin Bailey says his aim is to "provide a set of actionable recommendations that help you ensure that your jobs don't slow down the print production workflow … without affecting the visual appearance that you're trying to achieve." This is an important goal because it is the challenge that printers everywhere face, every day. The print job must look as the buyer of the job intended. That's a tall order because the buyer of the print job may not be the designer, have control over how many layers are in the source file, or understand how the designer specified color. This book will help printers identify paths to optimize files, and will lay out a plan for educating clients about preparing files.

The creators of those files are also an important constituency. While designers seeking the approval of their clients will naturally work to create innovative content, the ability to print the file is most often not their focus. Instead, they finely tune image content with cropping and adjusting hues and saturation, pick spot colors,

© Global Graphics 2021. All rights reserved.

and create multi-layer designs, all vital topics but also ones where file preparation may be secondary to selling the product or service. This book will help designers understand the opportunities digital printing provides and will give insight into how to make their design tools produce the most printable files.

One of those insights concerns a common element of professional designs, the layers of transparency used to mask elements. No matter why transparency is used in the file, it is likely to have consequences for how fast the file will RIP. Blending and overprinting are also common design elements that impact RIP speed. This guide explains why transparency, blending and overprinting may behave differently across the range of Digital Front Ends.

Another element of this book covers the wide-ranging world of variable data. More than the simple process of inserting a name or address into an open space in a file, today's variable data opportunities range widely, from merging data directly from a database to using data to inform decisions about what color scheme, images, or logos will be used. With this book, both designers and production teams can see what happens as the result of the decisions they make on these and other issues.

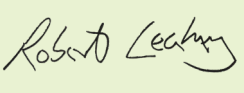
The final essential element in this book is the explanation of why poorly constructed PDF files hurt business. For example, the impact of specifying

"Your reward for studying and following the book's advice on this topic and others will be well-formed PDF files, ones that will process efficiently through the RIP and print just as the content owner intends."

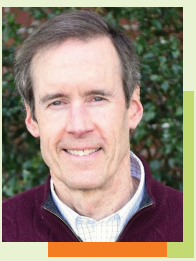
Bob Leahey and Pat McGrew,
KeyPoint Intelligence.

© Global Graphics 2021. All rights reserved.

spot colors when they aren't truly needed should be part of conversations with all designers. The same is true for how assets are positioned in a file. Simple design decisions can make a big difference at every stage in the print supply chain! And because the often hidden process in the printshop is the workflow, by following the guidance in Optimizing PDF for Variable Data Printing, you'll be able to ensure that you don't ask your workflow to do more than is necessary. Your reward for studying and following the book's advice on this topic and others will be well-formed PDF files, ones that will process efficiently through the RIP and print just as the content owner intends.



Bob Leahey
Director Keypoint Intelligence



Pat McGrew
Managing Director McGrewGroup, Inc

© Global Graphics 2021. All rights reserved.

01 | INTRODUCTION

Martin Bailey, Distinguished Technologist, Global Graphics Software

Digital printing is helpful either when it's used for a job that could have been run on a conventional press (e.g. offset or flexo) but does it better, or when it allows you to achieve something that a conventional press can't do and allows the brand or printing company to add new value.

In the first category digital printing usually wins in handling short print runs for any one of many reasons: it allows for a faster turn-round; it's more economical because no plates are required; it generates less waste; or it can eliminate warehousing when a larger job is called down in many smaller deliveries. In addition, it can reduce time to market for new designs. These are all important reasons for using digital print, but not the subject of this guide, which is specifically focused on using digital because it can do something that conventional presses cannot do: to make every copy of the print different.

Twenty-five years ago variable data printing was only significant in the transactional space (printing phone bills and credit card statements) and in coding and marking (printing serial numbers and best-before dates). That was partly because digital print technology was in its infancy; printing in double-digits of pages per minute at over 300dpi was only just coming in. It was also because, while computing power was arguably up to the task, most companies simply did not have the data to do much in the way of personalization.

In the commercial print space variable data print passed through "the next big thing" stage fifteen years ago, and direct mail printed fully digitally has dropped into my mailbox most days for quite a while. Alongside that it's also widely used for personalized catalogues, college enrolment packs, Christmas cards and photobooks … and many more items.

© Global Graphics 2021. All rights reserved.

More recently variable data print is marching onwards into new print sectors, into wide format, labels and packaging, and into some areas of industrial print for décor, textiles, product decoration, ID cards etc.

This huge variety of jobs is created and managed by an equally huge variety of software, from specialist composition tools to general purpose design applications carefully configured for VDP. The jobs are then consumed by workflows involving (or even completely within) the Digital Front End (DFE) for a digital production press, where jobs must be imposed, color managed, RIPped, screened etc. and printed in much the same way as a non-variable job.

Those tools are installed, configured and used by an equally broad variety of people. Obviously, there are brand owners and marketing campaign managers involved, but also graphic designers, experts in data management and manipulation, specialists in combining that data with the graphical design, and last (but definitely not least) the operators for digital production presses and of the finishing equipment that is used to cut, trim, and all the other finishing, conversion, mailing, delivery and fulfilment operations required for that specific job in that specific print sector.

That variety of use cases, creation software, print options, people and equipment involved means that a huge number of complex and sophisticated workflows have evolved over an amazingly short time. And most of the time they work just fine – the job is printed and delivered as designed.

Over the years several projects and publications have been developed to assist with designing jobs for digital print. One of the first, and probably the most influential and famous was "Design for Digital" from the Printing Industries Alliance Digital Print Council. Some colleges and universities also offer courses in how to design a job for digital print, although increasingly it seems that designing for print is not regarded as a key skill in graphic design courses. On the plus side, some production press vendors run courses for designers.

© Global Graphics 2021. All rights reserved.

Pretty much all of those publications and courses, however, are either "how-to" guides for getting the most out of a specific piece of composition software or are aimed at maximizing the response rate on direct mail pieces. Both of those are important goals, but they don't address the important question of how the internal structure of VDP jobs should be optimized for efficient processing.

And that's the goal of this guide: to provide a set of actionable recommendations that help you ensure that your jobs don't slow down the print production workflow ... without affecting the visual appearance that you're trying to achieve.

After all, as a brand owner, you don't want that phone call at 5pm on a Friday saying that your job will not be shipped that day after all because it ran too slowly on the press. And you don't want your print service provider telling you they will need to charge you more if your jobs continue to run more slowly than anyone else's.

Or, as a print service provider you don't want your press schedule disrupted by jobs that take significantly longer than they should have. The press should be running at full engine speed at least most of the time that you have press runs on your plan.

As a side benefit, several of the recommendations set out in this guide will also ensure that your PDF files can be delivered more efficiently on the web and to PDF readers on mobile devices in a cross-media publishing environment.

Some of the advice in this guide can be applied quickly and easily by a graphic designer, using their current tools. Some is intended more for the software companies building composition tools. If all of us work together we can greatly reduce the chance of that "heart-attack" job; the one that absolutely, positively must be shipped today … but that runs really slowly on the press.

Martin Bailey
Chief Technology Officer, Global Graphics Software

© Global Graphics 2021. All rights reserved.

02 | What is variable data printing?

For the purposes of this guide, "variable data printing" is any job where all the pages are related (i.e. you're not simply printing a huge number of separate jobs), but where many of the pages, taken as a whole, are unique.

That uniqueness may be in only a small way – maybe only a single graphic or line of text that's different to every other instance of the same underlying page design – while the majority of the artwork and content is the same.

Or the uniqueness may be very significant, e.g. when printing photo books, or designs where the graphics are all variations from a small number of 'seed' designs.

The variation may be for process control; shipping and returns; security; customer- or campaign-driven or for any other reason. Most of this guide (and all of the recommendations in chapter 12) is applicable for any of those.

But I don't print pages!

This guide will talk about 'pages', even though many print sectors don't print 'pages' of output, e.g. in labels, packaging, industrial print etc. We needed a word to refer to a collection of graphics that is printed together, and 'pages' seemed like the most intuitive word. It's also aligned with PDF terminology, which has 'page objects'.

If you're not printing pages, please simply mentally translate to 'label' or 'carton' or whatever else is appropriate for your use case.

In the same way, vocabularies differ depending on the class of work being printed, and on the background of the speaker. Somebody from a data center environment will often talk about a 'controller' driving a 'printer'; in the graphic arts or in labels & packaging it's more common to discuss a 'Digital Front End' (DFE) driving a 'digital press'. If we'd tried to be inclusive in the wording used throughout this document it would be twice as long, so we've settled on a relatively common set of terminology that we're confident you can translate from as necessary, into whatever is appropriate for your specific work.

© Global Graphics 2021. All rights reserved.

03 | Representative variable data print workflow

Variable data printing workflows vary significantly, depending on who is responsible for what parts of the whole process; and so does the terminology used to describe workflows and individual steps. It may be useful to bear the model shown below in mind while reading this guide as a way of translating between the terminology used here and the vocabulary that you typically use.

You will also see that the digital press has been subdivided into two components. The Digital Front End, or DFE, is where all the software processing of the job is applied. This means rendering from PDF into a raster (a collection of pixels), including color management and halftone screening. The resulting raster data is transmitted to the printing unit (inkjet heads or electrophotographic drum) to be marked onto the paper.

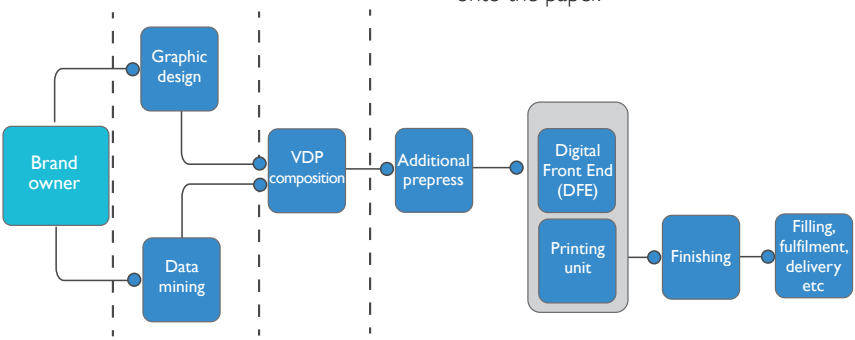


FIG 1. Simple example variable data creation and printing workflow.

In this diagram the dashed, vertical lines show several different points in the workflow at which jobs may be handed over from one department, division or company to another.

The distinction is important because the printing unit will have a known, fixed speed for some combination of number of colorants, resolution etc. But the conversion from PDF to raster requires far more complex calculations, and does not take a

© Global Graphics 2021. All rights reserved.

fixed time to process each PDF page, or square meter or sheet etc. The tooling and techniques used to create a PDF file can have a huge impact on how long it will take to render it.

That difference is really the reason for this guide. In order for a company running a digital production press to make a profit the printing unit must usually be running at close to rated "engine speed".

For variable data jobs that means that the DFE must be rendering those pages at engine speed. And that means that the graphic designer and VDP composition operator, amongst others, can have a direct impact on the profitability of the printing operation, and on pricing and other consequences.

"While maybe not as sexy as the choice of a new digital press, choosing the correct workflow to keep the machine running is ever more important. As print and packaging markets change, with ever shorter runs and quick turnaround demanded by customers who are not necessarily print professionals, it is critical that correct print-ready files are generated into the press queues with minimal manual involvement. In 2020 successful low-touch and even no-touch automation workflows are being offered that will examine and process PDF files uploaded by customers. This may include both production functions (pre-flight check and repair service, color management and imposition for the particular press) with increasingly administration features, of planning, costing, materials and delivery also involved in a seamless end-to-end system. Automating the production and business processes of printing, finishing and despatch allows print business models to change, satisfying the customer requirements in our increasingly connected world."

Sean Smyth
Analyst and Consultant

© Global Graphics 2021. All rights reserved.

04 | Where is variable data printing used?

Over the last few years the range of print sectors taking advantage of variable data printing has grown significantly. Just to give a few examples:

- Even in delivering print on paper, new areas have picked up variable data, e.g. because book publishers want a unique barcode on the cover of every copy of a book to assist in managing returns.

- The use of personalized URLs (PURLs), has grown out of direct mail, but can now be found on advertising at bus stops or hoardings. It allows brand owners to identify where somebody saw the ad and read the QR code for more info.

- Brand owners and regulators have increased requirements for the tracking of high-value goods and pharmaceuticals, leading to a growing industry to deliver full track and trace of products. Incorporating unique identifiers in the packaging of those products in a variety of overt and covert ways is an important part of that, and some of those processes is addressed by using variable data printing.

- An increase in selling a standard product with a custom, personalized label via the web. Coca Cola, Marmite, Nutella and many others have taken this route, both for marketing, and because a personalized product can be sold for a much higher margin than a standard product through a supermarket.

And in the near future:

- Digital printing can already be used for RFID antennae, but several vendors are getting close to being able to print the RFID 'chip' itself at production speeds on digital presses. At the moment RFID is usually used to record the product SKU, but would there be a benefit in RFID chips carrying unique identifiers, e.g. for track and trace in pharma?

- Digital printing of textiles for fashion is already starting to make use of techniques to avoid wastage, e.g. by only printing the fabric that will be used. What opportunities are there for unique clothing, either by varying the cut for a custom fit, or by using a unique variation of the pattern for every instance?

- And many more.

The real attraction of digital printing is that it opens so many doors!

© Global Graphics 2021. All rights reserved.

05 | Unique vs. personalized vs. versioned

As discussed earlier, different sources use a wide variety of terminology for variable data print. One area that can cause problems in discussing issues is in differentiating between different categories and use cases for variability. Some treat variable data, personalization, versioning etc. as equivalent, while others, mostly those aiming at marketing professionals, differentiate between a much larger number of categories.

The load on the DFE for the digital press, and the requirements for handling wastage and for delivery of the printed material, vary a lot with how the variable data is intended to be used. So, for the purposes of this guide, we're going to define a number of terms.

5.1 Versioned

Versioning is when several smaller print runs are produced instead of a single large run, and where the graphical content of the page varies between those runs. A common example would be a European brand creating packaging for a food product. Recent regulations on the minimum size of type used for the ingredients list means that it's no longer viable in many cases to produce a single pack with the ingredients list in all languages. So instead of, say, a print run of 5,000,000 copies, there may be a dozen different print runs between 100,000 and 1,000,000 each.

In many cases each version is still a reasonably substantial print run, so no special optimization is required in order to process it through the DFE on the press. In other words, this kind of versioning falls outside of the scope of this guide.

5.2 Traceable

A traceable job is when each instance of the printed item has sufficient information on it to be able to determine where and when it was printed, and by whom, perhaps by inclusion of a location code, batch code, date of print, filling or conversion etc.

Traceability does not introduce any special processing requirements either in printing or in onward delivery.

A normal level of wastage in printing, finishing or filling, etc. will not normally cause any problems.

© Global Graphics 2021. All rights reserved.

In many cases the variable data is very simple and will often be added using a coding and marking station somewhere in the production process, rather than being printed at the same time as the main graphic design.

FIG 2. Traceability can be enabled using unique, serialized identities on product packaging and labels

Coding and marking

For many years coding and marking has been regarded as a completely different operation from printing graphics. It focuses on very high speed, very reliable printing of simple graphics such as text and barcodes, to carry serial numbers, sell-by dates, process control or shipping information, and some elements of track and trace. It tends not to have such high-quality requirements as printing of the associated graphics, but it often needs to run consistently in environments that are somewhat hostile to delicate electro-mechanical systems such as inkjet printers.

This book is not designed to provide guidance for pure coding and marking systems, mainly because most of the time they're not consuming PDF graphics anyway; their graphical requirements do not necessarily warrant doing so. The data being printed is also often streamed, where data is still being generated when the first copies are printed; PDF is not designed for streaming.

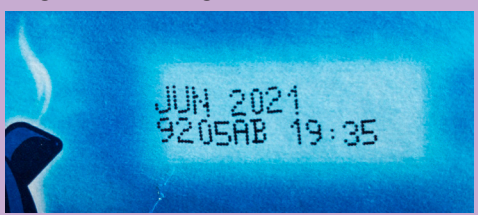


FIG 3. Typical low-resolution coding and marking print.

But the availability of fairly compact and reliable inkjet printers that run at suitable speeds for small filling lines is increasing. And that means that some products can use a single device for printing the graphics of a label (for instance) and items that might traditionally have been treated as coding and marking. This is definitely not something that could be deployed universally, but it does illustrate how different classes of work have a tendency to converge.

© Global Graphics 2021. All rights reserved.

5.3 Trackable

A trackable job is where the brand owner or other creator of the job knows in every instance where it has been delivered to. It almost always implies traceability, but also requires far more stringent supply chain management, and often means that every printed piece must be uniquely identified. Trackability means that the brand owner knows where the product is (or at least should be), whereas traceability allows the source of a product to be determined if that product is available to inspect.

For the purposes of this document we're going to assume that trackability includes a need for anti-counterfeit measures. The two are often closely linked because there's far more value in tracking a genuine package if you can also prove that it is genuine.

Some measures can be applied as part of the print process without variable data, by hiding information, or by using designs that are difficult to reproduce. Examples include encoding a static message into images (steganography), or copy-evident designs such as void pantographs, or complex and difficult-to-copy artwork such as microtext or guilloche (geometric lathe work).

Other measures will encode a traceable identifier for the print location, date, batch number etc., and these will use variable data techniques. Many products are now available to add a pattern of marks over all or part of the printed piece in a way that is not immediately obvious to the naked eye, but which can be detected by special tools, such as apps using the camera on a mobile phone.

FIG 4. Sophisticated marks can be used to carry unique information and to make counterfeiting more difficult.

Best practice in anti-counterfeit is to use a range of different overt and covert techniques, such as using a substrate that is chemically marked or contains security threads, potentially taggants in the ink etc.

The output of such tools varies significantly on how much additional load it places on processing in the DFE for a digital press.

© Global Graphics 2021. All rights reserved.

Trackability imposes many more requirements on process control at every step of printing, finishing and distribution. At the very least it requires waste to be identified so that the spoiled instances are also tracked. In some cases replacement instances must be printed and inserted into the appropriate sequence of product.

5.4 Personalized

A personalized job is where each instance of the printed piece was created for a specific recipient.

Personalization can be used in transactional prints and in push marketing where a brand owner has sufficient information about the intended recipient. The most obvious example of this is in direct marketing / direct mail.

Outside of marketing, however, personalization is normally only applicable in a pull model, where the product has been specifically requested by or for the recipient. If you're selling a jar of Nutella with an individual name on the label you must have a way to get that jar to them. So, this is a model best handled through a web portal where people order jars as gifts.

Caution: *The definition of personalization used in this guide does not match how the word is used in some print sectors. In labels and packaging in particular 'personalization' is often used as a synonym for versioning.*

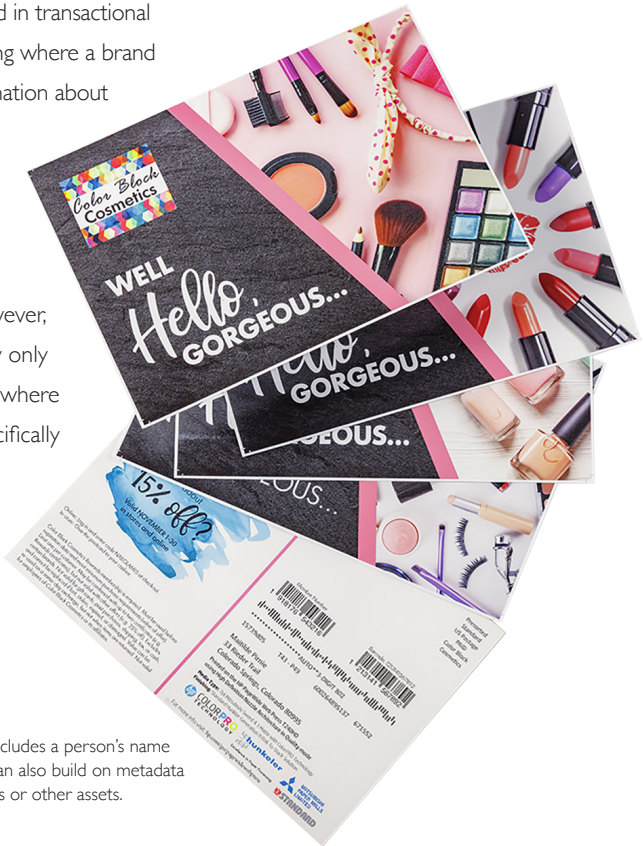


FIG 5. Personalization usually includes a person's name and often their address, but it can also build on metadata held by a brand to select images or other assets.

© Global Graphics 2021. All rights reserved.

Personalization implies all of the same requirements on process control as for traceability. In much of the transactional space there are legal requirements that personalized documents such as credit card statements must be sent to the right person and at the right time. But in any situation where the recipient is expecting the printed piece (e.g. the jar with their label on it) there is an obvious obligation to ensure that it is created and delivered.

It's worth noting that the majority of bottles sold in the famous "Share a Coke" campaign by Coca Cola did not use personalization by this definition.

In each country a number of the most common first names was selected, and a large number of labels were printed with those names (more or less) randomly mixed on them. In a sense it was closer to a versioned job than a personalized one. Only the individual bottles available through roadshows and the web site were truly personalized … because only in those situations did Coca Cola have a route to deliver the bottles to the correct recipient.

FIG 6. A form of personalization can be used even in a 'push' delivery model; the "Share a Coke" campaign simply used common first names in each geographical area.

© Global Graphics 2021. All rights reserved.

5.5 Graphically unique

Graphically unique jobs are sometimes described as "Every Page Is Different", or EPID. Used in this guide it means that each instance of the printed piece is different, but where that difference is not functionally significant.

The archetypal graphically unique jobs (outside of photobooks) are those created using HP's Mosaic and Collage software or by Hybrid Software's PatchWorker technology in Packz, which place copies of one or more seed graphics on each printed piece, but varies how they are positioned, scaled and rotated. The result is that a designer can create a campaign where every label, carton, postcard, brochure, phone case or item of clothing has a unique design on it … without having to generate every single one of those variations by hand.

But when processed through the DFE for the digital press, each of those unique designs must be rendered individually. There's no opportunity to apply clever techniques to optimize away some of those processing requirements, as described in chapter *10. Technologies for printing variable data.*

FIG 7. New solutions allow every piece to be unique, while retaining a brand feel based on aspects of the design or color selection.

© Global Graphics 2021. All rights reserved.

5.6 Personalized and graphically unique

An extreme example of personalization is photo finishing and photo book printing. It's personalized to the extent that an individual has ordered something specific and unique.

The fact that it's rare for more than one or two copies of each job to be printed doesn't preclude it from being treated as a "variable data" print job, because photo books, personalized calendars etc. are usually batched up into long sequences of books for different recipients, within a single submission to the print stream.

But it's also graphically unique, in that every page is different.

This, then, is one of the hardest cases to handle. It carries all the responsibilities of excellent process control that any personalized job requires, but also places the highest load on the press DFE because every page is different … and handling unique images creates one of the greatest demands. In other words, it's a range of products that will benefit even more than most from the optimizations described in this guide.

FIG 8. Photo books are the archetypal example of graphically complex print jobs where only one copy is usually printed.

© Global Graphics 2021. All rights reserved.

06 | Not just full color

To start with variable data print was just in black (or black and red for bank statements!). The transactional world is a few years into a transition to color workflows. But even "full color" isn't the limit for variable data.

If something can be processed digitally, with every instance different, it can be incorporated into VDP. And that opens up opportunities in premium VDP, with premium margins.

A few examples to think about include:

- 'Full color' usually means CMYK, but variable data can also be printed using an extended-gamut ink set, usually using two or three of Orange, Green and Violet inks as well as CMYK. This is often especially appealing in labels and packaging, where a wide gamut press can often match more brand colors accurately.

- Variable data in white ink (as opposed to a simple knock-out from CMYK as is common for print on paper) or varnish. On some presses this will also enable texture or tactile effects. If your press can't achieve those then it can probably be combined with an embellishment device that can.

- Unusual colors such as pastel, fluorescent or metallic inks. And 'Invisible' inks, such as those that become visible under UV, either as part of an anti-counterfeit program, or just for fun.

FIG 9. HP fluorescent inks image from drupa 2016.

© Global Graphics 2021. All rights reserved.

• Variable data in cold foiling for a higher gloss finish than many metallic inks, and also for surface effects.

FIG 10. DIGITAL METAL® from KURZ: digital embellishment solutions for enhancing fine graphic artwork with customized shiny eye-catchers.

• Variable braille; if your press or embellishment engine can achieve tactile effects with varnish it may also be able to print braille.

Care must be taken to ensure that the appropriate standards of dot size and height can be met.

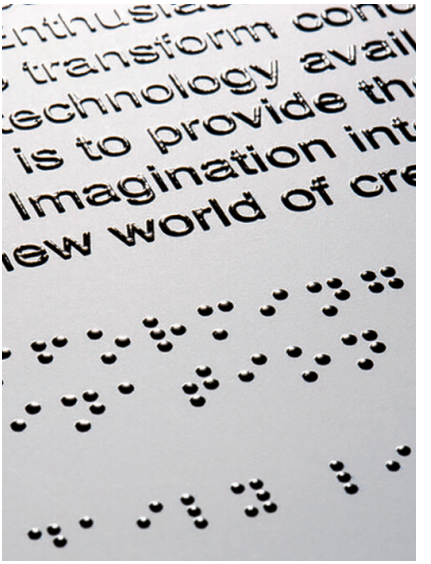


FIG 11. Braille printed with inkjet. Courtesy of Roland DG Mid Europe Italia.

• Variable data in cut-outs using laser cutters. At present complex cuts tend to make the cutter slower than your press, and so may need to be processed off-line if you don't wish to slow the press down. But they are becoming faster every year.

© Global Graphics 2021. All rights reserved.

For most of these opportunities the additional data just needs to be provided as spot colors in a PDF file sent to print. For wide-gamut color, brand colors should be specified as spot colors as well, to allow the emulation on the press to be adjusted for an optimal color match (*but* see *1.2.4.3*).

If the color and embellishment elements of the design are to be handled on separate devices, and those devices are not tightly integrated, you'll need to take a bit more care to ensure that the color print and the embellishment are matched correctly. At the very least you should include registration marks to allow them to be aligned properly; on web-fed devices eye-marks will also be useful.

Some embellishment engines can also read barcodes off the color print to identify the job, page or recipient number etc. and to make sure that the right variable data is added to the right color print. Ask your print representative what you should do to include such information, if anything.

"Today's graphics business is about delivering high-quality data to the page as fast as possible. It's about marrying efficient data processing with superb output quality-control in a fully automated print workflow. This new Global Graphics guide to optimizing PDF for variable data printing offers recommendations to maximize variable data delivery, whilst optimizing digital press performance, including output quality and speed."

Laurel Brunner
Consultant and trade journalist.

© Global Graphics 2021. All rights reserved.

07 | What file formats are used for delivering variable data jobs?

Even just five years ago this would have been quite a long section, because there were a lot of different file formats used for variable data printing. But that field has been significantly simplified in the last decade, to the point where only two file formats still have any significant usage for variable data, especially for new jobs. Legacy jobs being adapted for new revisions may still use a legacy format.

In the pure transactional space, and for jobs with multiple hundreds of thousands or millions of recipients in direct mail, AFP/IPDS is still a significant player.

The regulatory framework for approvals for designs of financial documents is largely built around AFP, especially in North America, and that is very slow to change.

For everything else, from most direct mail that hasn't come from the banking and financial sector, through commercial to labels, packaging, wide format and industrial print, PDF has emerged as the format of choice.

There are other formats still in use, such as PPML and IJPDS, but not at anything like the same scale as these two, and most commonly for legacy applications.

There are several subset standards built on top of PDF for production printing, including PDF/X and PDF/VT. See *Appendix B – Relevant standards.*

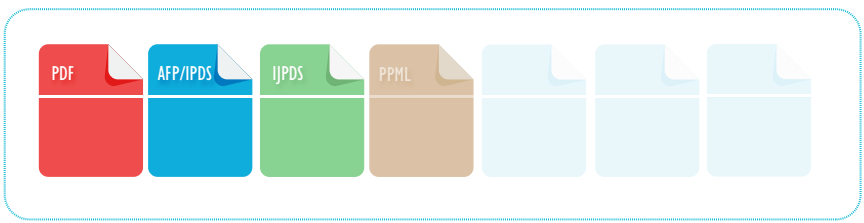


FIG 12. Formats other than PDF and AFP/IPDS are used primarily for legacy applications.

© Global Graphics 2021. All rights reserved.

08 | Why does optimization of VDP jobs matter?

If you're printing work that doesn't make use of variable data on a digital press, you're probably producing short runs. That's the most often cited reason to use a digital press instead of an offset or flexo press, although there are others. But "short runs" very rarely means a single copy.

Let's assume that you're printing, for example, 50 copies of a series of booklets, or of an imposed form of labels. In this case the DFE on your digital press only needs to RIP each PDF page once and then print the same raster 50 times.

Let's assume that you're printing on a press that can produce 100 pages per minute (or the equivalent area for labels etc.). If all your jobs are 50 copies long, you therefore need to RIP jobs at only two pages per minute (100ppm/50 copies). Once a job is fully RIPped and the copies are running on press you have plenty of time to get the next job prepared before the current one clears the press. Obviously longer run lengths give the RIP even more time.

But VDP jobs place additional demands on the processing power available in a DFE because most pages are different to every other page and must therefore each be RIPped separately. If you're printing at 100 pages per minute the DFE must RIP at 100 pages per minute; fifty times faster than it needed to process for fifty copies of a static job.

" Today, digital printing is mostly used for short runs, which is usually not interesting from an economic point of view. A print job of 10 copies requires often the same administrative effort as a print job of 10,000 copies. Printing variable data enables higher print runs and therefore more profit. Mastering variable data is essential to exploit the potential of digital printing. In the past, the definition of variable data was dominated by proprietary formats of individual manufacturers. Open standards such as PDF/VT allow a wider use of variable data for everyone. However, there is still a lot of training to be done. "

Stephan Jaeggi, PrePress Consulting.

© Global Graphics 2021. All rights reserved.

Each minor inefficiency in a VDP job will often only add between a few milliseconds and a second or two to the processing of each page, but those times need to be multiplied up by the number of pages in the job. An individual delay of half a second on every page of a 10,000-page job adds up to around an hour and a half for the whole job. For a really big job of a million pages it only takes an extra tenth of a second per page to add 24 hours to the total processing time.

If you're printing at 120ppm the DFE must process each page in an average of half a second or less to keep up with the press. The fastest continuous feed inkjet presses at the time of writing are capable of printing an area equivalent to over 13,000 pages per minute, which means each page must be processed in just over 4ms. It doesn't take much of a slow-down to start impacting throughput.

If you're involved in this kind of calculation you may find the digital press data rate calculator at https://blog.globalgraphics.com/choosing-the-class-of-your-raster-image-processor-rip/ useful:

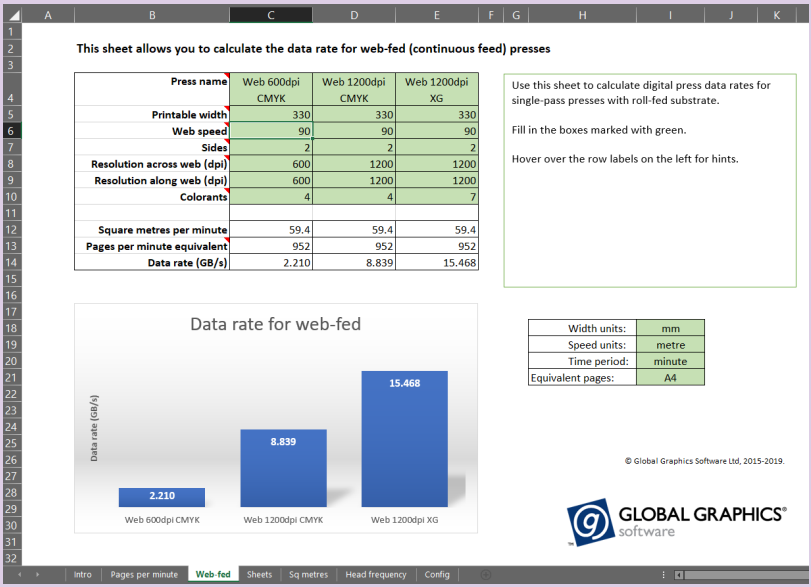


FIG 13. Global Graphics' digital press data rate calculator.

© Global Graphics 2021. All rights reserved.

This extra load has led DFE builders to develop a variety of optimizations. Most of these work by reducing the amount of data that must be RIPped and are described in section *11 - High level view of VDP optimizations*: RIP once, use many times. But even with those optimizations a complex VDP job typically requires significantly more processing power than a 'static' job where every copy is the same.

The amount of processing required to prepare a PDF file for print in a DFE can vary hugely without affecting the visual appearance of the printed result, depending on how it is constructed.

Poorly constructed PDF files can therefore impact a print service provider in one or both of two ways:

- Output is not achieved at engine speed, reducing return on investment (ROI) because fewer jobs can be produced per shift. In extreme cases when printing on a continuous feed (web-fed) press a failure to deliver rasters for printing fast enough can also lead to media wastage and may confuse in-line or near-line finishing.

- In order to compensate for jobs that take longer to process in the DFE, press vendors often provide more hardware to expand the processing capability, increasing the bill of materials, and therefore the capital cost of the DFE.

Once the press is installed and running the production manager will usually calculate and tune their understanding of how many jobs of what type can be printed in a shift. Customer services representatives work to ensure that customer expectations are set appropriately, and the company falls into a regular pattern. Most jobs are quoted on an acceptable turn-round time and delivered on schedule.

> But occasionally a customer supplies a file that takes much longer than expected to process and disrupts the whole schedule.

Depending on how many presses the print site has, and how they are connected to one or more DFEs this may lead to a press sitting idle, waiting for pages to print. It may also delay other jobs in the queue or mean that they must be moved to a different press. Moving jobs at the last minute may not be easy if the presses available are not identical. Different presses may require different print streams or imposition and there may be limitations on stock availability, etc.

© Global Graphics 2021. All rights reserved.

Many jobs have tight deadlines on delivery schedules; they may need to be ready for a specific time, with penalties for late delivery, or the potential for reduced return for the marketing department behind a direct mail campaign. Brand owners may be ordering labels or cartons on a just in time (JIT) plan, and there may be consequences for late delivery ranging from an annoyed customer to penalty clauses being invoked.

Those problems for the print service provider percolate upstream to brand owners and other groups commissioning digital print. Producing an inefficiently constructed PDF file will increase the risk that your job will not be delivered by the expected time.

This guide is designed to help you avoid making jobs that disrupt and delay the printing process, increasing the probability of everyone involved in delivering the printed piece hitting their deadlines reliably, and achieving their goals effectively.

You shouldn't take these recommendations as suggesting that the DFE on any press is inadequate. Think of it as the equivalent of a suggestion that you should not fill your brand-new Ferrari with cheap and inferior fuel!

“ Personalization in marketing communication and packaging is growing rapidly to support demographic and population segment demands. Whether it is for the sake of targeted marketing or security, understanding how to use variable data printing to support these requirements is critical. This guide is a great primer for anyone who needs to design and produce to these new requirements and breaks down what can be done and how to do it. ”

Dave Zwang, Zwang.com.

© Global Graphics 2021. All rights reserved.

09 | Who is responsible for optimizations?

Over the years a very wide variety of different approaches to composition for VDP have emerged. To over-simplify the market somewhat, most of these can be characterized as taking various positions along a scale between two points:

- 'Simple': Complete 'backgrounds' for variable data jobs are created by a graphic designer. The composition tool can import those backgrounds as assets, often as PDF pages, which can only be placed and positioned as atomic units. Relatively simple variable text and graphics, often including barcodes, can then be placed, usually over the top of the background.
- 'Complex': A wide variety of assets can be imported into the composition tool as individual images, logos and graphics. Both variable text and other data and a variety of rules can be added. Rules may select which assets should be placed for each recipient of the printed piece and where those should be positioned. In some cases variable data can be used to construct graphics on the fly, such as pie charts for financial reports, or personalized images.

Many tools can be used in a way that allows the operator to select where their particular usage falls on this spectrum. As an example, even when a tool capable of a

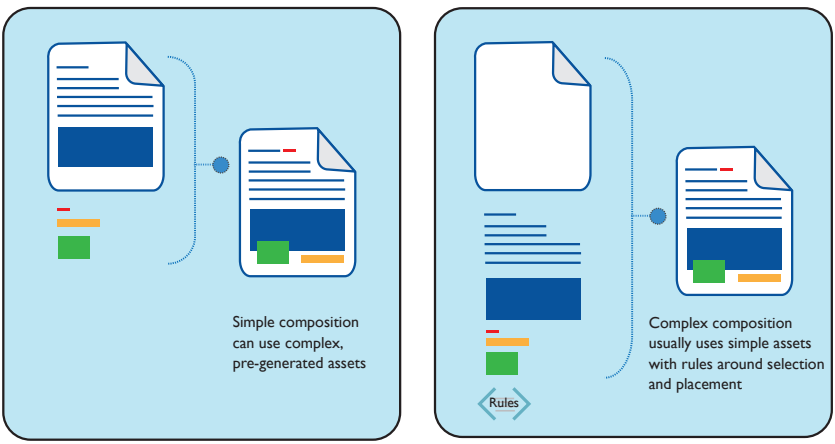


FIG 14. Example assets for constructing similar final pages in simple and complex composition engines.

© Global Graphics 2021. All rights reserved.

'complex' design process is used, it's often possible to place an asset that happens to be a full-page PDF file which already contains multiple and potentially complex graphics.

This variation in tools and in how they can be used implies an equal variation in how all the people involved in creating a job can affect the efficiency of the resulting PDF file, and therefore where the responsibility for doing so will fall.

If a composition tool is used at the 'simple' end of the spectrum all of the choices about the use of transparency, image resolution, font embedding etc. are likely to have been made by the graphic designer when creating the 'background' assets. The composition tool will quite often include the PDF objects from the asset into the PDF file for printing as-is, without making any changes other than those required to reference them from the composited pages. In this scenario the graphic designer is largely responsible for the efficiency of the result.

At the other end of the scale, a composition tool may be used in 'complex' mode, where all assets are supplied as single images and relatively simple graphics; the rules defined in the composition engine manage most of the PDF print file construction. Responsibility for how efficiently the job prints in this case is shared between the designer, the developers at the composition tool vendor and the operator of that tool, depending on the richness of the rule set that can be used.

If the rule set allows the operator to make significant decisions around the use of transparency, for instance, those decisions must be made wisely. On the other hand, the tool itself will usually make decisions about if, and how, fonts should be embedded. If image down-sampling is available it may be configurable by the user or be applied automatically by the tool.

Composition tool vendors may also provide some relatively simple pre-flight feedback to identify use of assets or decisions by the operator that might reduce efficiency at the print site.

© Global Graphics 2021. All rights reserved.

10 | Technologies for printing variable data

When variable data printing first started, it was often achieved by printing the variable data digitally on top of a "pre-printed shell" that had been produced using an offset or flexo press. This technique is still used in some cases for simple variable data because it can be cost effective. But it obviously limits the data that can be used, because nothing that has already been printed on the shell can be changed. So it tends to be restricted to use only variable data in black, often placed in areas that were deliberately left unmarked on the shell.

Over the last decade the transactional and direct mail sectors have migrated strongly away from pre-printed shells towards what is commonly described as a "white paper workflow" or "White Paper Factory", because all of the graphics, including variable data, are printed in a single pass onto the substrate. Obviously, outside of transactional and direct mail the term "white paper" may not be accurate, but we'll use it in this guide because there are no widely used alternatives.

A white paper workflow has advantages over pre-printed shells in that it places no constraints on the variability that can be achieved. One example that is often cited is the ability to place variable text in white over a background tint or image … without needing a white ink. But that flexibility does come at the price of increasing the workload for the DFE on the digital press, which is precisely why this guide has been created!

FIG 15. HP Indigo 6K.

© Global Graphics 2021. All rights reserved.

Between the two extremes of pre-printed shells and white paper workflows are a couple of other options:

An imprinter may be mounted on a conventional (offset, flexo etc.) press, or on the finishing or converting line, or even on an inkjet web press to add the variable data. In a sense this is a bit like a pre-printed shell workflow, but without a requirement for warehousing the shells between printing them and adding the variable data. It's most often used for simple graphics, a step or two more complex, and with slightly higher quality requirements than coding and marking.

And finally, a growth market at the moment, especially in labels and packaging, is for hybrid presses. These use a sequence of print stations, which may use different printing technologies, most often flexo and digital, but sometimes also including screen printing for whites and coatings. Like imprinting, the concept is very close to pre-printed shells without the need for warehousing. But hybrid presses vary in a couple of ways because they are often used to apply key items on a label or carton. This means the quality requirements for the variable, digital print are higher, and that they often use color-managed full color digital print instead of just a black overlay. The combination means that designing for hybrid presses requires some unique skills and experience to get the best out of the system.

FIG 16. Cut sheet print module for imprinting from HP Specialty Printing Systems.

FIG 17. The Elan 500 HD press from Delphax Solutions.

© Global Graphics 2021. All rights reserved.

The optimizations described in this guide are intended primarily for use in white paper workflows but are also likely to provide some benefit where the variable graphics printed on the digital station of a hybrid press are reasonably complex.

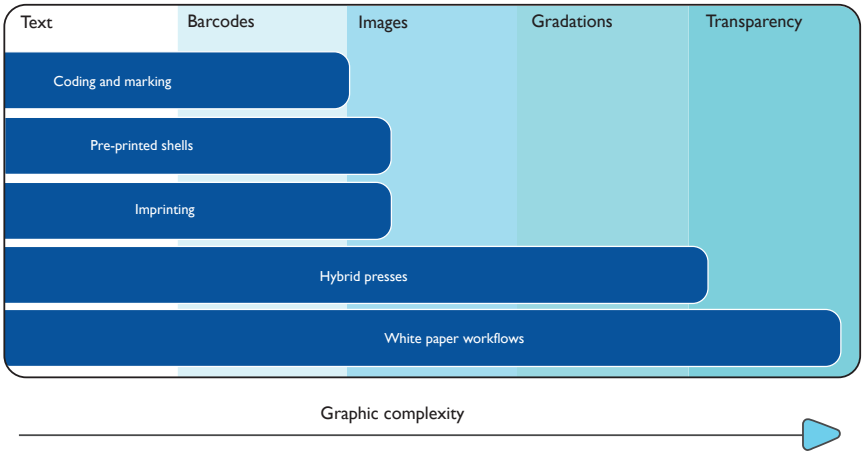


FIG 18. Simplified view of common levels of graphical complexity in different variable data print environments.

The graphics often used for pre-printed shells and imprinters will often be sufficiently simple that there isn't much that will really gain from optimization, but if you're pushing the limits of those technologies this guide may help!

© Global Graphics 2021. All rights reserved.

11 | High level view of VDP optimizations: RIP once, use many times

A very short run of a non-variable job on a digital press tends to mean that you're producing at least a few dozen copies. In other words, each page is processed once in the DFE (color managed, RIPped, halftone screened etc.), and then sent multiple times to the press. The DFE doesn't need to process pages at the same speed that the press engine can print them.

But if you're printing a variable data job it's likely that many pages will be unique; most pages will be at least slightly different to every other page.

Obviously this is not a universal rule; if you're printing invoices, for example, it's common for the back of every sheet to be the same as the back of every other … but even in that case there may be an invoice number or date added onto the back of the sheet.

Building a DFE to be able to process whole pages as fast as the engine could consume them when every page is different is relatively expensive, so the DFEs for many digital production presses include optimizations designed specifically to reduce the processing requirements for VDP jobs.

When a VDP piece is designed a variety of assets of various forms are collected together. Some assets are intended to be used multiple times, while others are associated with a single recipient or personalization. Both single- and multiple-use assets may include images, graphics (e.g. barcodes, maps, logos etc) and, of course, text.

All of the assets are placed and positioned according to a set of rules. Those rules might be as simple as a mail merge in Microsoft Word, where placeholders are included in a template for the document, and then replaced with text from a separate data file. In more sophisticated environments additional information from a database about each recipient is used to select from the assets available.

© Global Graphics 2021. All rights reserved.

Thus 'gold' members of an organization may see one version of an asset, while 'silver' or 'basic' members see different ones.

A classic direct marketing example is a mailer sent out to people who have previously bought a particular make of car a couple of years after that purchase to invite them to come in and view this year's model. Each piece might include a photograph of a car of the same class as the one they purchased and perhaps in the same color. Thus, if they had bought a sedan, they'd see an image of this year's sedan, if they bought a sports car, they'd see a sports car.

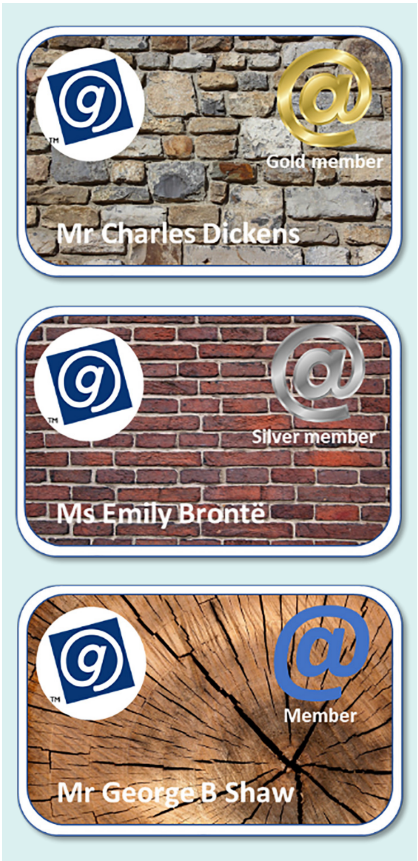


FIG 19: Different categories of recipient will often see different assets.

In addition there might be a map to the dealer that they bought from last time, the name and contact details for an appropriate sales representative, etc.

All of the assets required to reproduce the pages are then included in the PDF file and sent to be printed. The PDF can be viewed in any PDF reader and would display as a series of fully laid out pages. It could be processed through a DFE in that way as well … but often not at high enough speed to keep the press running at full engine speed.

The optimization process in a DFE is usually more or less the opposite of how the composition engines built the print stream in the first place. The PDF file is examined to identify graphics that are used as a group multiple times.

© Global Graphics 2021. All rights reserved.

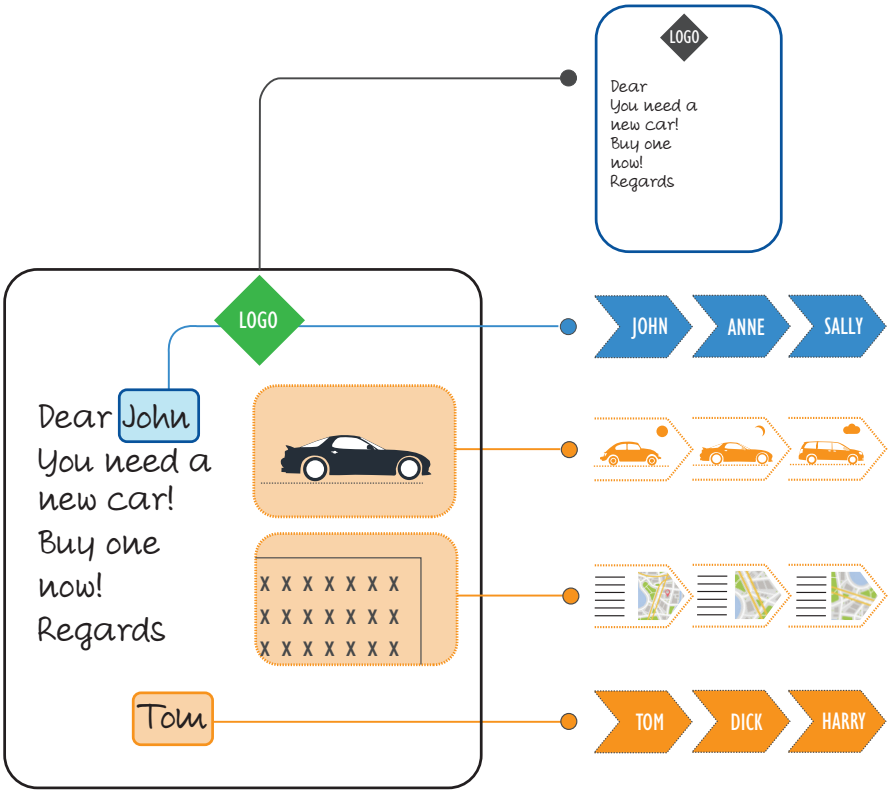
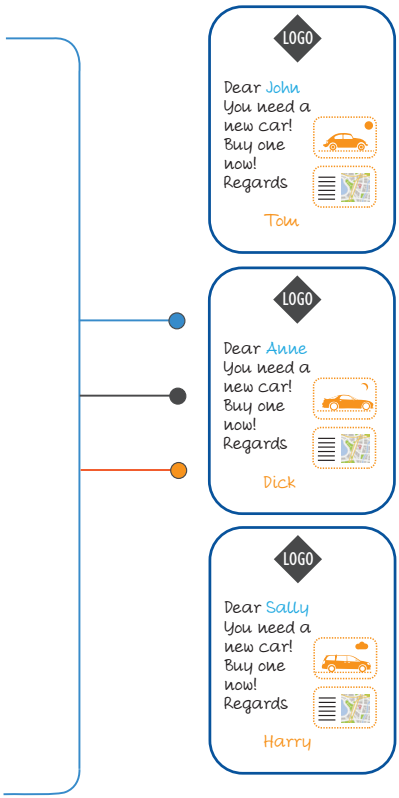


FIG 20: A simplified view of VDP optimization in the DFE

Those are then processed separately and stored, along with data recording where they were seen in the job. Those elements of each page that are only used once are also processed. Processing here typically means applying all color management and rendering to a raster format at the same resolution, and using the same colorants, as the digital press.

Finally, the various components that make up each page of the job are re-composited building a raster for each page from the partial-page rasters for each component.. This can be done in software, on a GPU, or on the driver board for the digital press itself; but it's usually better to apply any halftone screening after recomposition to avoid screening artifacts at the edge of each graphical component.

© Global Graphics 2021. All rights reserved.



WHICH ARE LATER RE-COMPOSITED INTO PRINTED PAGES

While this may seem a very complex method it means that many of the original assets from the design phase are only processed once in the DFE. In a well-designed and implemented system the overhead for splitting the job apart and then re-compositing the rendered components again is a lot less than the time saved by not having to RIP each asset in the job hundreds, thousands or even millions of times.

Different technologies from various vendors handle each step of this process in different ways. Some, for example identify re-used elements by looking at the number of references to PDF constructs called XObjects, while others also review all graphics and identify sequences that are repeated irrespective of how they are structured into objects within the file.

XObjects are effectively graphical assets defined within the PDF file itself. Each one holds an image, or a collection of graphics such as text, rules and color fills in a way that means they can be re-used many times in the same job without having to include the data multiple times. XObjects can be nested into complex combinations of graphics and/or images together.

Once re-used elements are identified, some systems coalesce them together by determining which collections of elements are used together on multiple pages with consistent positioning relative to each other. This minimizes the number of components required to construct every final page in the job. Achieving this coalescing automatically, flexibly and intelligently has a huge impact on the overall throughput of the DFE and is a key distinguishing factor between RIPs and DFEs from different vendors.

© Global Graphics 2021. All rights reserved.

12 | Making efficient PDF files

In every print workflow one rule overrides virtually everything else: the printed result must be what the person signing the check wanted and expected. This guide is not intended to restrict the ability of marketing departments and graphic designers to achieve the desired visual appearance of printed work. It provides guidance on easing the path to the most efficient production of that design … whatever that desired result might be.

This section sets out guidelines for avoiding tripping up the print production workflow with your PDF files for VDP. At the highest level almost all of them boil down to a very simple maxim:

> Don't ask the print workflow to do more work than necessary if that doesn't change the look of the printed result.

There are often many ways of achieving the same visual appearance which can vary significantly in the amount of processing required to print them. Sometimes the most efficient method for the print company requires a little more work for the origination company, and sometimes there's a win-win where improved print performance can be gained by making a few changes that also result in a PDF file that can be shared more efficiently on the web and on mobile devices.

The effect of much of the advice below, such as using images at an optimal resolution or discarding cropped image pixels, will vary significantly depending on how the graphics in question are used in the job. Optimizing an image that is used in exactly the same way on the output for every recipient of the job will have a very minor impact, because a well-designed DFE will only process that image a few times (possibly only once) and re-use the results multiple times.

On the other hand, optimizing images that are personal to every recipient (e.g. images custom-built to include the recipient's name) can have a huge effect because those images must be processed many times, once for every single recipient. Graphics that are used for some subset of the recipients, usually based on some metadata about the recipient, fall somewhere in between.

© Global Graphics 2021. All rights reserved.

If you only have the time to focus on parts of your workflow you should concentrate on the graphics that are individual to each recipient.

Most of these recommendations are relevant to the designers and composition operators in the trenches. A few are so deeply into the technical details of constructing a PDF file that they can only really be addressed by the developers who create and maintain the VDP workflow software that we all depend on. Those few have been split out to a separate section at the end.

“Like no other file format, PDF/VT creates the basis for a native data exchange within the age of mass customization in print, for all of the participants in the workflow. PDF/VT is the guarantee for a predictable end product - without having to generate gigantic amounts of data - for the creator of a file as well as for the processor. This makes PDF/VT the enabler of a new era in print.”

Bernd Zipper, CEO Zipcon Consulting and president of Initiative Online Print.

Where do these recommendations come from?

Global Graphics' Harlequin® RIP has been used in DFEs for a variety of different digital production presses for nearly two decades. Over that time we've received a lot of sample files from our OEM partners, primarily the press vendors themselves. If a job is too slow we often get to see the file and analyze why it's reducing throughput.

Sometimes we respond by optimizing the processing of specific constructs within a PDF file, and that's driven our development of the fastest RIP in the world.

Sometimes it's been more appropriate to point out to a customer that including thousands of copies of a 20-megapixel image on a single page, each at 6000ppi, may not be a very efficient way of making a file.

And through all that analysis we've continually learned what causes problems. In this guide we've turned those problems round to provide recommendations, all triggered by real world examples of jobs that didn't follow them!

© Global Graphics 2021. All rights reserved.

12.1 Optimizing images

As a general rule images tend to take longer than vector graphics and text to process in a DFE. A photographic image will often use quite a large number of different colors, each of which must be appropriately color managed. In addition there is simply more data involved which must often be copied between memory locations, and the difference between the effective resolution of the source image and the resolution of the output device must be resolved.

These operations only take a few milliseconds individually, but multiplied over all the images in a job they can add up to a significant total.

At the same time images are commonly re-used within a VDP job; they may form part of a static page background, or a small number of images may be selected from, each being used for a proportion of the recipients (like the car images in the example in 11, *High level view of VDP optimizations:* RIP once, use many times. Thus being able to process each of a relatively small number of images only once, and then re-use the result many times can significantly increase the throughput of the DFE.

It's worth noting that many of these recommendations around image handling will also make a PDF file more appropriate for multi-channel delivery, e.g. by the web, email or to a mobile device because they will reduce the file size and allow a more resource-constrained viewer to display them correctly.

12.1.1 Set photographic image resolutions appropriately

There's a general rule of thumb in conventional print that you shouldn't place photographic images with an effective resolution greater than double the halftone screen frequency that you're using, because you won't gain any quality from going higher. So if you're screening for an offset press at 150lpi (lines per inch), for instance, images should normally be included at, or just under 300ppi (pixels per inch).

The most appropriate image resolution for digital presses varies somewhat for each one, depending on the printing heads, media and screening used, but aiming at around 300ppi is still a pretty good target for most. Using an effective image resolution for photographic images higher than, or even close to, the output resolution of the press is virtually never productive.

© Global Graphics 2021. All rights reserved.

The image content can also affect this slightly; a soft and dreamy image can sometimes be placed at a significantly lower resolution, while one with high-contrast fine detail may benefit from a slightly higher one. To play safe in an automated workflow you may choose to select a resolution that is enough to maximize quality for the sharpest and most detailed images, say 350ppi or, for best results, ask your digital press vendor what they recommend.

As you can see from the box below it can be very easy to use an image at several times the required resolution. In the example the image is at 1000ppi on the page, about three times what is required. That trebling applies in both the height and width of the image, so there are actually nine times as many pixels as necessary, which can significantly impact performance in the DFE. Just imagine what would happen if the same image file had been placed at only 1.5 x 2 inches (3.75 x 5 cm); there would then be 36 (9x2x2) times as much data as required.

Effective image resolution

When an image is placed onto a page the original resolution of that image is largely irrelevant; what matters is how many pixels there are per inch on the final printed page. As an example, if you have a photograph from a 12MP compact camera it'll probably be approximately 3000 pixels by 4000 pixels. If that's placed on the page as 3 inches by 4 inches (7.5 x 10cm) the effective resolution is about 1000ppi (4000/4). That would usually be about three times as much as you need in each dimension.

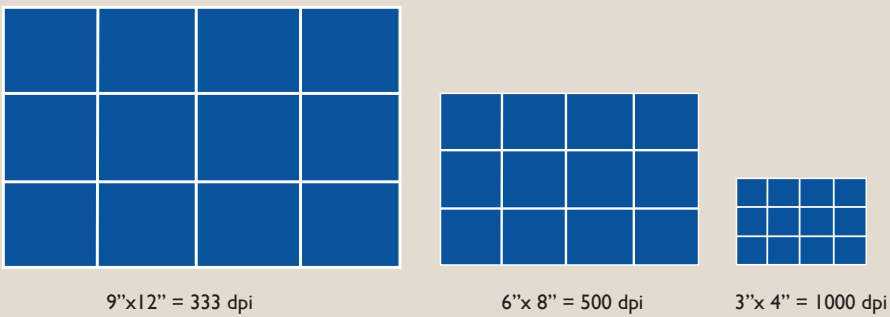


FIG 21: The same 12-megapixel (4000x3000px) image placed on the page at three different sizes.

If the image is cropped the calculation must be made from the uncropped dimensions.

© Global Graphics 2021. All rights reserved.

A variety of tools are available for optimizing image resolution, and some composition tools can also do this automatically.

Note that this section applies only to photographic images (where each pixel may represent one of a number of tone values for each colorant) including both color and grayscale. Copy-dot scans, screen grabs and other synthetic images usually benefit from higher effective resolutions, with the optimal value normally being at the same resolution as the press itself. But watch out for moiré between the original image resolution and the press resolution if you don't match them exactly.

12.1.2 Discard cropped pixels from images

If an image is heavily cropped the portions outside the cropped area should be completely discarded rather than simply hidden using a clipping path. Even though the clipped-out pixels won't typically be color managed etc., they will typically still need to be read from the PDF file and decompressed in order to find the pixels that are actually required.

Cropping images can sometimes be efficiently combined with a resolution reduction step.

FIG 22: Discard cropped pixels from images.

© Global Graphics 2021. All rights reserved.

12.1.3 Optimizing personalized images

Some asset creation or composition tools can create images that are personalized for each recipient of a VDP piece, by writing their name in creative ways within the image itself, for instance. In most cases the proportion of the image that carries the personalization is quite small. It is often more efficient for the whole image, without personalization, to be included once for all recipients, with a smaller image (or images) overlaid in the correct position to carry the personalized area. This means that the un-personalized image can be treated as static data and processed once even though it appears on many pages. The personalized image(s) will be treated as variable data and processed for every recipient … but being much smaller that processing won't take as long.

Of course, it's vital that the small, personalized, image(s) are exactly aligned with the whole background image and set to use exactly the same halftones to avoid any artifacts along their edges. Placing the personalized image as a masked image so that none of the pixels of the background image are duplicated will help to avoid artifacts if the alignment is not 100% correct. If you can't guarantee achieving that it's probably better not to try this optimization and to accept that the job will process a little more slowly.

As an alternative, you can sometimes achieve a very similar effect much more simply and easily by using a specialist font, without needing to generate personalized images at all!

12.1.4 Avoid image interpolation

The PDF specification includes a flag that can be included in an image to instruct the DFE to interpolate or up-sample the image.

Interpolation is a relatively slow process and should be avoided if possible. If a photograph is used at such a size that it does not achieve the minimum image resolution appropriate for your press it should be up-sampled during or before the creation of the PDF. Ideally you may wish to consider using a different image or cropping it less tightly to ensure that you achieve a high-quality print. If neither can be done the image should be included as-is, without requesting interpolation; the image quality is unlikely to be noticeably different from an interpolated one.

Again, this recommendation is most relevant for photographic images. There are cases where pre-rendered images of assets such as text, especially saved at one bit per pixel, can benefit from image interpolation, most notably when the effective image resolution is higher than the printing resolution.

© Global Graphics 2021. All rights reserved.

12.1.5 Don't use images for flat colors

Occasionally we see files where somebody has very carefully created an image where all the pixels are the same color and placed that in the design instead of just drawing a vector rectangle and filling that with color. On its own this is usually more likely to cause confusion when a job is preflighted than it is to slow the DFE, because that image often ends up at a much lower effective resolution than the preflight profile is set up to allow. But if it's combined with transparency it will also tend to slow the job processing, so it's best avoided.

The only real exception to this recommendation is if you're deliberately trying to make a flat fill that uses exactly the same color as some pixels in an image. It's not that unusual to use different color management for images and vector fills, e.g. a photometric rendering intent for images, because that will yield the best, natural looking photographs, and a colorimetric rendering intent for vector colors in order to match brand colors as accurately as possible. Using a vector fill may not achieve the required color match in such cases.

12.2 Optimizing transparency

The very rich and flexible support for live transparency in PDF is an incredibly useful aspect of the format and is one of the reasons for selecting PDF over other page description languages for production print.

On the other hand, compositing transparent regions in a PDF file is much more processor intensive than handling opaque areas of a page.

As an example, consider two overlapping RGB images, both tagged with an ICC profile for ECI RGB in a PDF file.

When outputting to a digital press printing in CMYK with no live transparency involved the color of each pixel in each image must be transformed into tone values for CMYK, usually using ICC profiles. In most DFEs the results of the calculation for each set of RGB values from the image will be cached and re-used when another pixel using exactly the same RGB values is processed.

There's a reasonable amount of calculation involved, but nothing too heavyweight.

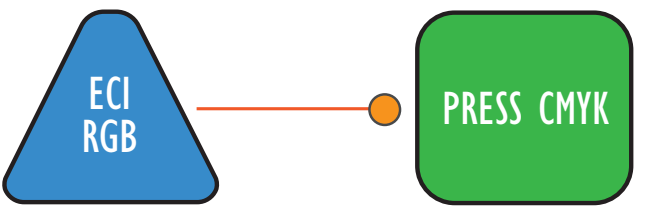


FIG 23: Color transformations without transparency are relatively simple.

© Global Graphics 2021. All rights reserved.

Now consider the same example where the two images are within a "transparency group" in the PDF file. In most cases that group will have a color space associated with it called the "blending color space", and in most cases that blending space will be sRGB, if only because that's the default in many design applications.

In addition a "blend mode" will be set. The blend modes allowed in PDF match those shown in Adobe® Photoshop®, including commonly used modes such as 'Normal', 'Overlay' and 'Multiply' and more specialized ones such as 'Soft Light' and 'Saturation'.

The colors of each pixel now need to be transformed from the source RGB (ECI RGB) to the blend color space (sRGB). Once in the blend space the two images need to be composited together according to the blending mode.

It's unlikely that the pixels of the two images are exactly aligned, so this composition means that the number of apparent pixels in the area where they overlap will increase.

And finally the resulting colors in sRGB must be transformed to the output CMYK of the press.

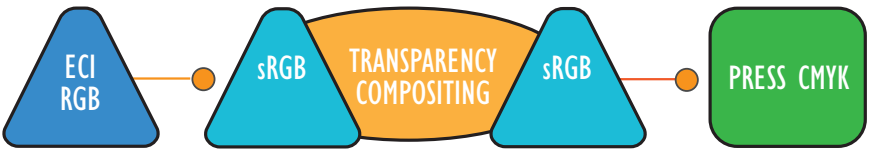


FIG 24: Color transformations with transparency requires significantly more processing.

As you can see this process at least doubles the amount of effort required in color transformations, even without taking into account the work to perform the transparency blending itself, which is significant for some of the blend modes.

It's fairly easy to 'accidentally' create transparency groups if you're placing PDF pages as assets into a larger area. The composition tool will probably place each

© Global Graphics 2021. All rights reserved.

original page as a Form XObject, and carry over information from the original page, or from a PDF/X output intent in the source document, by making that XObject into a transparency group.

The impact of using transparency in a VDP job depends on whether it's used in a 'background' graphic that's used many times on many pages, or if it's in variable data or a re-used object that overlays variable data. If it's in the background the VDP optimizations in many solutions will mean that it only needs to be processed once, which resolves the transparency. The result of that processing can be re-used multiple times so the extra work required in processing doesn't add much to the total job time.

If it's used in variable data, or in an object that overlays variable data, the VDP optimizations in many DFEs will be circumvented and the whole of the page may need to be processed as it stands without being able to re-use some or all previously processed elements.

The bottom line on transparency is that it's very valuable, but if it's not in the static background to pages it should be avoided, as long as that can be done without changing the final printed appearance.

12.2.1 Don't flatten transparency

It may seem strange after the previous section to say that transparency shouldn't be flattened. But flattening transparency upstream of the DFE can have two significant unwanted effects:

- The transparency effect can sometimes be replaced with a huge number of very small graphics in order to try to maintain exactly the same visual appearance. This not only bloats the file size, but it can make the job even slower to RIP than working from the live transparency would.

- If the flattening is not performed with a detailed knowledge of the resolution and other capabilities of the press the job will be output on, it can introduce some unpleasant artifacts in the output, such as jaggies. It may also not make use of any extended gamut inks (Orange, Green, Violet etc.) or special spot inks (White, Varnish, Foiling, etc.). Even if you do know the full details for the press that will be used, a pre-flattened job would be harder to transfer to another press at the last minute if you needed to.

12.2.2 Avoid invisible transparency effects

The most common usage for live transparency in PDF is for drop shadows, but even that use should be avoided if it doesn't result in an effect that's visible on

© Global Graphics 2021. All rights reserved.

the final printed piece. For example, do not include drop shadows on images that are printed on a black background unless the shadow will also fall on another element where it will be visible, such as another image on the page.

FIG 25: This image has a drop shadow on it, but it's completely lost against the black background..

Clearly there are exceptions to this where the drop shadow would still be visible on a print, even if it isn't on a computer monitor, such as where the drop shadow paints in a rich black (e.g. black plus 40% cyan) and the background is printed with only black ink.

In the same way, if all you're doing is adding drop shadows to text or images that fall entirely on a white background, you don't need to use transparency at all; a simple shading pattern will do everything that you need. Of course, if any of the graphics with drop shadows overlap each other you will need to use transparency, so that the shadows fall across the elements behind correctly.

If assets are being created in off-the-shelf design tools and then integrated with variable elements in a composition tool it may be difficult to avoid the use of transparency in drop shadows, because many design tools offer a simple switch to

© Global Graphics 2021. All rights reserved.

add a drop shadow, which includes turning on the transparency. On the other hand, if everything is created and laid out within the composition tool it should be very achievable.

12.2.3 Use overprinting instead of transparency for black text and rules

Printers using offset lithography, flexography and other conventional print technologies have used a little trick to avoid registration errors between small black text and fine rules running over other graphics on a page for many years: they set the black elements to overprint. This means that the text and rules don't knock out of the other graphics, which means that you'll never see any white outlines as a result of misregistration. More recently we've seen a few instances where people have used transparency instead, using Overlay or Darken blend modes.

The potential for objectionable artifacts when using either approach is disappearingly small. The only visible effect likely is that the black won't be pure, but may have varying amounts of cyan, magenta and yellow added by graphics behind it. If these techniques are used only for small black text and rules it's hard to see that variation at all, even with a lens.

Where overprinting and transparency do differ, however, is in the speed at which the DFE can process them. A simple black overprint will often be very significantly faster, especially if the background behind the black elements is complex or includes high-resolution images.

12.2.4 Use clips rather than masks

Clipping an image, either to a smaller rectangle or to a more complex shape, can be done in several ways, and these vary greatly in efficiency:

a) A vector clip-path is by far the most efficient and should be used wherever possible.

b) If the creation workflow is such that a vector clip-path cannot be applied, use a masked image (a hard mask, encoded as an image with a Mask entry).

c) By far the most expensive in processing power is a soft mask (SMask), which is the only one of the three approaches that uses live transparency. These should only be used where a soft blend is required, e.g. between an image and a special effect frame.

Some applications use a soft mask to clip an image only because a hard mask at the same resolution as the main image would result in visible stepping around the edge. A vector clipping path will yield a smoother edge than most hard masks and would be a suitable alternative to a soft mask in many cases.

© Global Graphics 2021. All rights reserved.

When a special effect frame is added to an image it is usually printed on top of the image. It is far more efficient to reveal the real image through the frame using one of the following techniques than to add a soft mask to a frame supplied as an image:

a) Draw the frame using vector objects (far easier for some visual effects than for others). In this case nothing extra is required to reveal the image through the center of the frame.

b) Apply a clipping path to the frame object.

c) Use a masked image (with a Mask entry) rather than an image with a SMask entry.

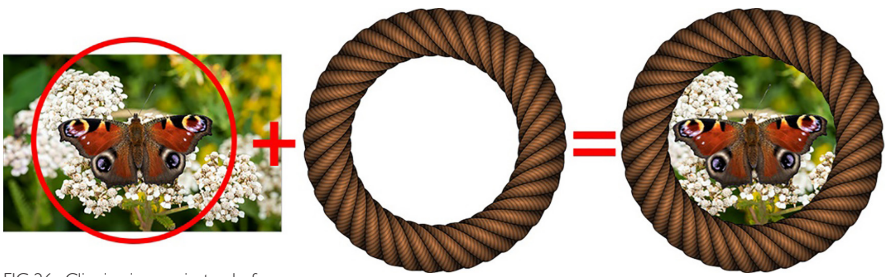


FIG 26: Clipping images instead of using transparency.

When using a frame with a complex irregular or non-rectangular shape that requires portions of the real image to be hidden so that they are not visible outside the frame, a clipping path should be used on the main image data as well. This often requires only a relatively rough outline as the clipping path only needs to fall somewhere in the area covered by the frame and does not need to track its edge exactly.

Of course, ideally the original butterfly image in FIG 25 should also be clipped to the minimum rectangle required to cover the inside of the frame, as described in *12.1.2.*

12.2.5 Pre-composite images with soft masks

Some VDP designs include the placement of one image with a soft mask over another background image, perhaps to achieve a soft transition from one to another. If it is possible to composite the two images with the soft mask into a single image before delivery to the DFE, the work required in the DFE will be greatly reduced.

There is little benefit to be gained from compositing multiple images without masks simply because they fall on the same page or because they overlap each other. The coalescing step of the VDP optimization (see *11*) will normally achieve this stage quite efficiently.

© Global Graphics 2021. All rights reserved.

12.2.6 Avoid using transparency for image ghosting

One effect that is sometimes used when placing a text block on top of an image is to 'ghost' the image behind the text, reducing its contrast and making it lighter so that the text can be read more easily. This can be achieved by placing a transparent rectangle over the image and behind the text, but that will mean that processing in the DFE will be very inefficient because it needs to resolve the live transparency. Either of these two techniques would more efficient:

a) If every use of the image requires the same size and position of ghosted area then the image and the ghosted area should be pre-composited, resulting in a single image and no transparency in the PDF

b) If the size of the ghosted area must vary for different recipients (e.g. because their address is printed in that space, and addresses differ in the number of lines) then it is better to include two copies of the image data, once for the full background, and once for the ghosted area. The image used for the ghosting may be pre-adjusted before inclusion in the PDF, or the adjustment may be applied using a transfer function.

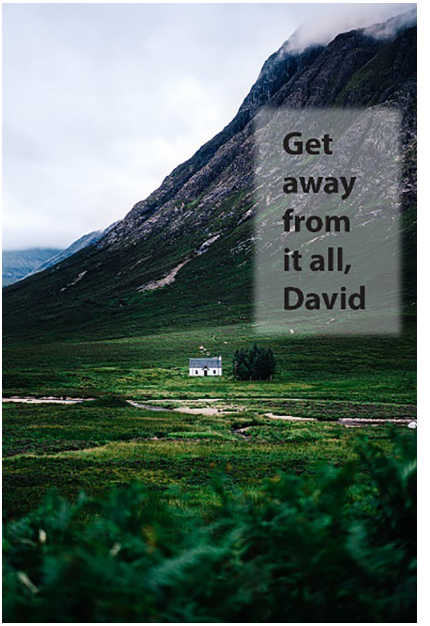


FIG 27: Ghosting images to allow text on top of them to be read.

For the maximum performance gain, the parts of the image that never fall within the ghosted area may be discarded in the second copy of the image, rather than just clipped out. Care must be taken to ensure that the two images are exactly aligned in that case. Even though this technique increases the amount of image processing required, it can increase overall performance because image processing is much faster than transparency compositing.

© Global Graphics 2021. All rights reserved.

“ Digital printing is finding a home in many areas of package printing. So far the greatest focus has been for short run and “drop in” orders that are problematic for converters to manage. Digital presses have been deployed to address these problem areas and along with them have come enhanced workflows that help ease the path of work through all levels of the plant operations. The next level of adoption for some packaging will be for variable data... every package being unique. Moving to this level of implementation will require even more automation and integration of hardware, software and processes all through the packaging supply chain. The complexity of such implementations will require a great deal of cooperation and tools that aid in that cooperation will be extremely valuable. I commend Global Graphics for having the foresight and initiative to develop such a tool and for having Martin Bailey author it. ”

Kevin Karstedt, CEO, Karstedt Partners LLC.

12.2.7 Avoid unnecessary color space conversions for transparency

As mentioned above, a transparency group in the PDF file can have a blending color space defined within it. In these cases the colors of graphics within the group must be transformed from their original color space into the blending color space, and then subsequently into the output device color space.

Many PDF files have transparency groups with a blending color space set to sRGB, simply because that’s the default in a number of mainstream design tools, while the output device color space for print is usually CMYK (or some variant upon that).The key message here is that the transparency won’t add any additional transformation of the color information if the blend color space of the group matches either the source color space of all graphics within the group or the device

© Global Graphics 2021. All rights reserved.

color space. The transforms may occur at a slightly different place during processing, but the same amount of transformation is required.

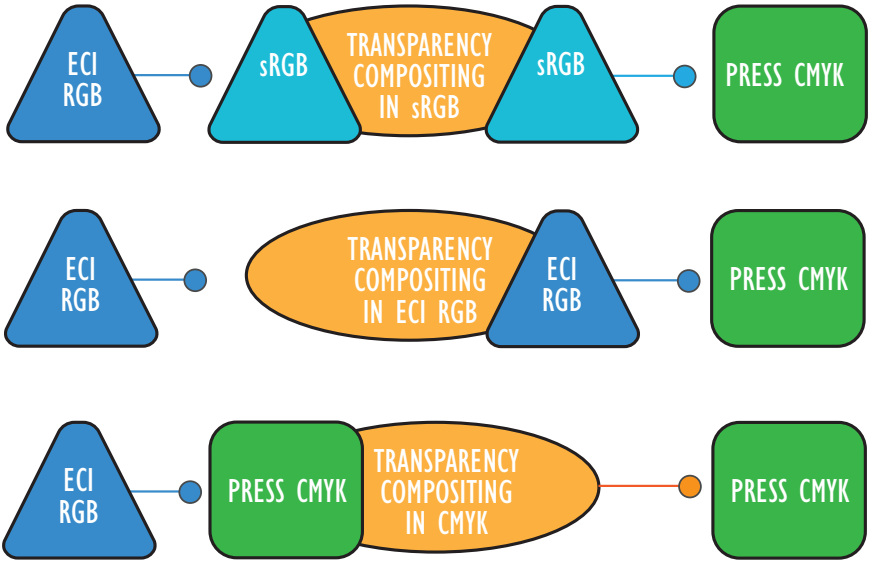


FIG 28: Choosing the blend color space carefully can greatly reduce color transformations required.

But if the blending color space doesn't match either the source color space or the device color space, the colors of all graphics must be transformed twice instead of once, increasing the overall processing time.

If you can ensure that all graphics (especially images) within a group have the same source color space as the blending space, or, even better, the blending space matches the output device color space, then throughput in the DFE will be higher.

Switching the blending color space, especially between RGB and CMYK spaces, will often change the final printed color. If you're going to change the blend color space from something like sRGB to the output CMYK for maximum DFE performance, you need to make that decision early in your design process

© Global Graphics 2021. All rights reserved.

and ensure that the resulting output is approved. If you need to stay with blending in RGB you should ensure that the blend color space matches the source color space of as many of your images as possible (or vice versa).

Occasionally transparency group operations may be chained together if a group is defined within another group, although that is relatively rare. There can be good reasons for using this kind of construct in a non-variable job for commercial print, publication or newsprint work. One common example is when multiple PDF/X files that use different ICC profiles in their output intents are placed or imposed together; this might arise if you're placing display ads, for instance. If that kind of situation occurs in a VDP print job, however, you would be advised to review the creation workflow and unify your asset design process further upstream to ensure consistent and predictable output. In general nesting transparency groups should be avoided for VDP.

12.3 Optimizing vector graphics

Vector graphics are relatively quick to process compared to images, which is why this section is so short.

12.3.1 Barcodes and QR Codes

QR Codes and other barcodes can be represented on a PDF page in several different ways, including as an image or using vector graphics. Barcodes can also be drawn with a barcode font, although doing so with 2D barcodes such as a QR Code can be a bit fiddly.

In terms of processing speed a barcode font is typically the most efficient, but sometimes limits the opportunity to compensate for edge growth to maximize readability. Using vector graphics is also usually pretty quick.

FIG 29: A QR Code using 2x2 pixels for each module.

© Global Graphics 2021. All rights reserved.

The speed of processing a barcode recorded using an image (or image mask) depends very much on the image resolution used to encode it. If each barcode module is represented as a single image pixel it can be just as efficient to process as a vector representation … but that may trigger warnings or errors from preflight software if the effective resolution of the barcode image is below the threshold that has been set. Increasing the image resolution to avoid preflight warnings by using multiple pixels per module will tend to result in slower processing, but not normally enough to worry about too much.

Composition vendors can assist with barcode quality (readability) if they're using vector graphics by turning on automatic stroke adjustment for bar codes (using the SA graphics state parameter in the PDF) to minimize issues if the scaling is not absolutely correct.

12.3.2 Avoid unnecessary smooth shades

Smooth shades were added into the PDF specification in the late nineties, and provide a way of defining a variety of graduated tints or vignettes. They can be very useful but tend to take a little longer than a simple flat fill to process, especially if they happen to interact with any transparent graphics on the page.

Don't use a smooth shade where the final color doesn't vary across the object; just use a flat tint instead.

12.4 Optimizing colors

It's common practice to use spot colors when defining a color that will be used many times in the same design, especially when those are brand colors and will be used in multiple jobs. But using too many, or using them in sub-optimal ways, can impact the processing speed on the DFE.

12.4.1 Avoid overprinting spot colors

There are situations where spot colors really must be set to overprint, such as when you need a white ink under the color inks because you're printing on transparent, metallic or off-white substrates. Best practice in that case is to add that white on top of all other graphics in the design, and to set it to overprint. That way the other graphics can't accidentally knock out of the white. The object order in the PDF file itself doesn't have to have a direct relationship with the printing order, of course.

FIG 30: Example of flexible packaging using a white ink under the colors.

© Global Graphics 2021. All rights reserved.

And you'd usually do the same for a varnish; place it on top of all other graphics in the PDF and set it to overprint.

But if you're using spot colors for the color, most of the time you've chosen that color because that's what you want to see; you don't want it mixing with other colors. And if that's the case, make sure that you don't set it to overprint. On a digital press spot colors will very often be emulated using CMYK inks (or CMYKOG etc.) and managing overprinted spot colors so that they will emulate not just the color of the solid spot, but also the color resulting from overprinting with graphical elements underneath them, takes more processing, and therefore more time, not to mention being very specific to the press and inks that are being emulated.

12.4.2 Merge equivalent spot colors

When a PDF file is constructed with assets from a variety of sources you can sometimes end up with multiple spot colors that will print the same, but that have different names. As an example, you might have 'PANTONE REFLEX BLUE C' and 'PANTONE Reflex Blue'. The DFE won't merge those unless it's been specifically configured to do so, but if you merge them upstream in the design or composition stage you can make the job more efficient.

This same rule applies to spot colors that will be printed with real inks such as white or varnish just as much as it does to spots that will be emulated. Your jobs will run faster if you don't have two inks called 'White' and 'Tinta blanca', for instance.

In some cases you may genuinely need multiple hits of a spot, perhaps to increase opacity of a white ink, to build extra height when using a varnish for a tactile finish, or because you need a bump plate on the Magenta to achieve the vibrant reds that your design brief calls for. If the second hit should only be applied in some of the places where the first is to be marked, you really do need two spots with different names. But if you want the two passes on press to be identical, we'd recommend that you talk to whoever will be doing the printing to understand if they really need two copies in the job, or if they can configure the press to print that ink twice from the same separation. If you just put in two spot colors that will both end up being mapped to the same ink … you probably won't get a double-hit anyway!

© Global Graphics 2021. All rights reserved.

12.4.3 Did you really need thirty spot colors?

Some designers appear to define spot colors for every different color in a job, and we occasionally see PDF files with well over 30 spot colors in them. There's no problem in rendering those jobs (Harlequin RIPs can process PDF files containing up to 8,000 different spot colors), but it does increase memory requirements and processing load and therefore makes preparing the jobs for print slower.

A lot of design software will allow creatives to define swatches of colors either as process (using a mixture of CMYK), or as spot colors; you don't need to make something a spot in order to define it and to re-use it consistently across multiple graphics and multiple jobs.

Some spots will map directly to real inks on press (primarily white, varnish and other special effects), and you should use spots in the design stage for these. For other colors the main advantage of using a spot color is that it's possible in many DFEs to redefine the mix of process inks that will be used to emulate it on a digital press, and therefore to tune that color for the specific substrate etc. That's obviously important for brand colors, but much less so for other colors in the design.

So if you find the number of spot colors in your jobs creeping up into double figures we'd recommend that you take a look and see if all of them need to be spots, or if a process color in your swatch book would work just as well.

At the time of writing, Adobe Acrobat will only show the first 27 spot colors in a PDF file when using Output Preview. If you go over that limit it can slow your job down in other ways when somebody in the print workflow panics because it's not showing an important brand color!

12.5 Optimizing VDP layouts

As mentioned above (see 11) the ability to coalesce multiple graphics together to reduce the number of components that need to be re-composed together to form a final page can have a very significant impact on the throughput of the DFE. The coalescing process typically requires that multiple graphics must all appear on a significant number of pages together, and with exactly the same positions relative to each other in order to be grouped together into a single component.

© Global Graphics 2021. All rights reserved.

Some RIPs have the capability to adjust the drawing order of the assets and other graphics placed on the page, that is the order in which they are to be placed, with some behind or in front of others. Being able to re-order graphics allows them to be coalesced into groups even if they are not adjacent to each other in the drawing order. Of course, those solutions place great importance on avoiding any changes to the visual appearance of the printed page as a result.

Most of the recommendations in this section are aimed at maximizing the efficiency of the coalescing process so that fewer components are required to construct every final page.

12.5.1 Place graphics in consistent locations whenever possible

If you're creating several related page layouts that use the same assets (e.g. images) you may be able to generate each one by copying the previous one and making the necessary changes, or you may need to build each from scratch , depending on the composition tool you're using. In either case you can improve the efficiency with which the final job passes through some DFEs by ensuring that there are no unintended changes to the position or size of each asset on the page as you do so.

If a group of graphics are used together many times, then it's even more important

" The use of variable graphics and personalization adds complexity to the composition process. Designers may not be aware that efficient preparation of variable imagery and text for proper composition is critical to optimized RIP production and print quality downstream. This guide outlines a creative workflow process which creative designers should consider before designing with variable content. "

Mary Schilling, Co-founder Inkjet Insight and Co-author of The Designer's Guide to Inkjet

© Global Graphics 2021. All rights reserved.

to ensure that their positions relative to each other are consistent than it is for a single graphic. Varying their relationship will mean that the DFE's ability to optimize the job by coalescing them together as a group will be reduced.

FIG 31: These four pages illustrate different instances of a document where all of the blue 'text' and the blue 'image' (marked 'AAA') are intended to be used together, while the red 'text' and the green 'image' are unique to that instance. On the second page the green image is in a different location to those on pages 1, 3 and 4; but that will only matter if it uses transparency (or spot colors set to overprint that will be emulated using the process inks) as it's unique data and won't be shared between pages. The blue image on page 3 is also in a different position to the blue images on the other pages. Here it does matter, because it prevents the DFE from treating the combination of the blue text and blue image as being used together in exactly the same way on every page. The incorrect positioning may mean that the blue image is processed twice: once for pages 1, 2 and 4 (in combination with the blue text), and again for page 3 on its own.

The same comments go for placing all copies of a graphic or group of graphics at the same size and rotation.

If you have a good reason to move things around on the page then go ahead, but finding that the throughput of the DFE is reduced because you accidentally didn't place them in exactly the same position would be frustrating!

© Global Graphics 2021. All rights reserved.

Unfortunately this does mean exactly the same; a fraction of a point difference will often be enough to subvert any optimizations; after all, at 1200dpi each point is nearly 17 device pixels!

In the same way, some composition engines offer the capability to 'flex' layouts, to move some assets in response to differing sizes of something like a text block because some recipients have longer names or addresses, or the length of a list of items varies. Again, if that produces the exact visual result that you're looking for go ahead and use the option.

But if flexing the layout doesn't provide a benefit for you in the design or readability, turn it off and allow the job to process a bit faster at the print stage.

Consistent placement of graphics also extends to consistently placing clipping paths around repeats of the same set of graphics. If you're imposing something like postcards or labels, make sure that you use the same clip for all of them, even if that means the clip on some may extend off the edge of the media.

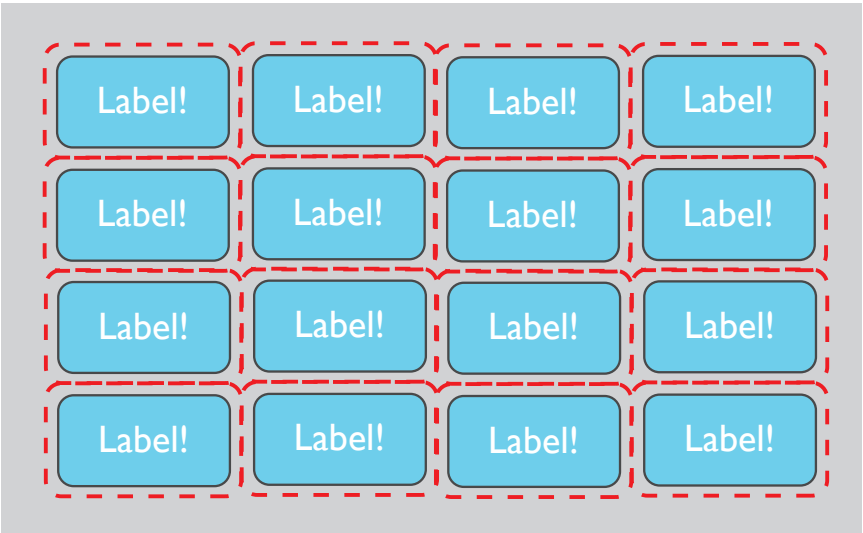


FIG 32: The gray area in this diagram represents the substrate in a narrow-web label press, with four lanes of labels imposed on. Each label station in the PDF file had a clipping path around it, shown in the diagram as a dashed red box. If the clipping path for each station is exactly the same with respect to the graphical content then the job can probably be processed faster in the DFE, even if that means that the clipping paths extend outside of the PDF page. The same would apply for imposed postcards etc.

© Global Graphics 2021. All rights reserved.

12.5.2 Avoid interleaving static and variable elements on a page

Many VDP designs boil down to a static 'background' that is used exactly the same on many pages, with variable data laid over the top of it, varying by the recipient of that instance, or a serial number or whatever. The variable data may be specific to a recipient (e.g. their name and address). Some may also be "semi-variable", where metadata about the recipient is used to select from a relatively small set of options (e.g. a logo for membership level, a map to their nearest store location, etc.).

The coalescing process in the DFE (see 11) will typically work best if it can merge all of the assets and other graphics for the 'background' into one or a small number of components to be re-composited later. It may collect sets of semi-variable assets and elements together as well, if they are used together in a consistent way. To take the example given in *High level view of VDP optimizations: RIP once, use many times,* of a map to the recipient's nearest store, it may be that that map is always used with a logo and a text address for that specific store, and with a sales representative's image and telephone number.

It's common to see PDF files where the assets and graphics on a single page are drawn onto the page in a fairly arbitrary order, so that 'background' graphics are actually drawn quite late, after many of the variable and semi-variable graphics. This often makes no difference to the visual appearance as long as the graphics drawn later don't fall on top of those drawn earlier. But it does mean that the coalescing step must work harder and its rules, designed to ensure that the visual appearance is not compromised, mean that it may not be able to collect graphics into a small number of large components, typically reducing throughput.

If you can design your assets and layouts in such a way that static background elements are drawn first, followed by semi-variable graphics, and then those specific to the current recipient, the coalescing stage can often perform better. At a slightly more detailed level, it's often worth trying to make sure that an image and the key line for that image are next to each other in the drawing order

12.5.3 Minimize object overlaps

Sometimes it's not possible to design the assets and layouts to allow them to be drawn in an optimal order as described in the previous section. In this case it can be useful to avoid graphics overlapping previously drawn ones unless it's required for the design. If objects don't overlap at all the coalescing step will have a lot more

© Global Graphics 2021. All rights reserved.

freedom to change their position in the drawing order to optimize the creation of groups of graphics.

The same goes for imposing multiple instances of something together on a larger PDF page, whether that's direct mail, postcards, labels, or anything else. It's very common to construct the PDF by creating a single instance of static, semi-static and variable graphics, and then to impose or step and repeat that across the available area (often with each instance encoded as a Form XObject).

If the background of one instance overlaps with the (identical) background of another instance, then the DFE may decide that it cannot safely treat the two as two copies of the same graphic.

That's especially likely if they use any live PDF transparency, because the overlap needs to be processed as an overlap to generate the correct visual appearance of that transparency. And if they can't be treated as separate graphics, then the DFE must process much more data because it must render both of them … or tens, or hundreds of them, however many there are on that PDF page. Some DFEs can be configured to ignore small overlaps in such cases, but in general it's better to ensure that the multiple instances don't overlap in the imposition.

FIG 33: Overlapping label designs; the yellow lines indicate the extent of each copy.

© Global Graphics 2021. All rights reserved.

12.5.4 Nest 'forms' and images appropriately

While some DFEs coalesce graphics automatically (see 11), others require that the coalescing is guided entirely by how assets and other graphics have been written into Form and Image XObjects in the PDF file. If you're using one of these DFEs the throughput can be significantly increased if you use some care in creating your own compound assets before placing them in the composition tool. Unfortunately this can cross the lines of responsibility between graphic designers and composition tool operators, and can make late changes to the page layout, or customization for markets using multiple output sizes (e.g. both US Letter and A4 pages) more difficult.

In the same way, a composition vendor can optimize throughput in some cases by replicating the hierarchy of single-use and re-used graphics in a hierarchy of form XObjects.

12.5.5 Don't mix variable and static data in form XObjects

Pushing too many graphics too deep into the hierarchy of form Objects (12.5.4) risks undermining the recommendation to minimize object overlaps (12.5.3), because some DFEs will treat everything in a form as being a single object.

For the graphic designer or composition operator this means that graphics that are only used for a single recipient should not be bundled into the same asset as graphics that are used many times for multiple recipients.

Having said that, it's common good practice to use a form XObject for each one of multiple, imposed instances of a design, such as a postcard, label or carton, and those will obviously usually include both re-used and variable data. In these cases you should try to ensure that all of the re-used data within that instance XObject is stored in another XObject that is referenced from the instance as shown in FIG 34.

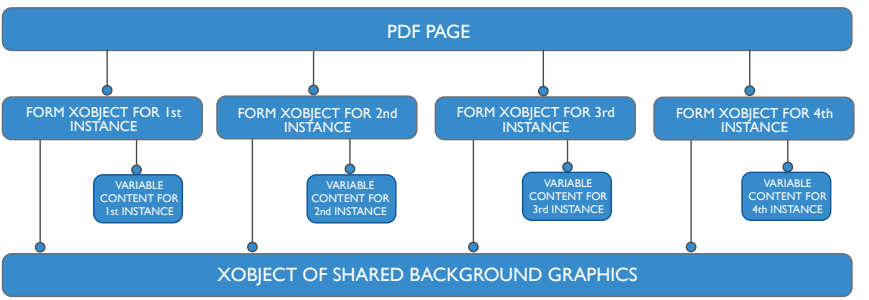


FIG 34: Tree of objects in typical variable data job where multiple instances of postcards, labels etc. are imposed together on each PDF page.

© Global Graphics 2021. All rights reserved.

12.5.6 Don't draw the same graphic multiple times

It may seem obvious that drawing the same graphics in exactly the same place on the same page multiple times may impact on performance, either directly or by reducing coalescing efficiency.

But it's something that we see quite often.

The same comment goes for drawing graphics and then hiding them completely with another graphic over the top. We've even seen cases where a complete page was drawn and then (we assume) the designer or composition operator decided to redo it, placed a white rectangle over what they'd done already to hide it and drew another complete page to replace it. The RIP will still need to do a reasonable amount of work to process the hidden first page, and it's just going to slow things down.

We recommend that you don't be that guy!

This doesn't apply to graphics drawn in PDF layers (optional content groups).
If a layer is hidden then it will have virtually no impact on processing time. So using layers for something like language versions and repeating graphic assets on multiple layers is fine.

12.5.7 Don't add unnecessary white fills

We've seen jobs where every page, or even every line of text, has a white rectangle placed behind it (as 0% gray or 0, 0, 0, 0 CMYK). Even though each one takes a tiny fraction of a second to process, every little helps, so don't add those rectangles if you have the ability to avoid doing so.

This obviously doesn't apply to adding background elements in white ink, e.g. for printing on transparent, metallic or off-white substrates. If you need the white there, add it!

12.6 Selecting and using workflow software

12.6.1 Evaluating security marking software

Pretty much every piece of software that you use to create a variable data PDF can have an impact on the speed at which your jobs will run on press. But one category that has come to our attention a couple of times is software that adds security overlays as a pattern of very fine dots over the top of the other graphics. If that's done using live PDF transparency then it may completely disable all variable data optimizations in the DFE, slowing everything down significantly. On the other hand, if it's simply done with an overprinted black it'll have virtually no impact on the performance of the DFE.

© Global Graphics 2021. All rights reserved.

Security marking software that amends raster objects or elements (e.g. using steganography) rather than adding an overlay has no detectable impact on processing speed for the final variable data PDF file, unless it means that the PDF ends up containing many more different images than it would have otherwise..

Very complex guilloches used for security and drawn as vector strokes can also slow processing, especially if each instance is unique.

If you're a brand owner, print service provider or converter planning to add security overprint software for anti-counterfeit or track and trace etc. we recommend that you include tests to determine if security marking will cause problems at print time in your evaluation process.

12.6.2 Play with the knobs

Many current composition tools have developed and evolved over many years, and they often contain controls designed to address historical issues in DFEs from various vendors. Not all combinations of options will work well.

If you're creating many jobs in one composition tool and sending them to the same printing workflow time after time then it's well worth reviewing the composition tool documentation, knowledgebase or forum to look for recommendations for optimal configuration for the DFE in that print workflow.

If that doesn't provide anything useful and if your jobs are processing slowly, you might simply try different settings in the composition tool to work out which ones help with the speed.

12.6.3 Managing jobs with many unique designs

Some variable data jobs are created using tools such as HP's Mosaic that uses one or more 'seed' graphics to generate a huge number of unique designs. Many of the recommendations in this guide apply to such jobs, especially those around images and transparency. Indeed, those recommendations are even more important for these jobs because the DFE is less able to cache parts of the job, and must process all, or nearly all, of every instance. Shaving a fraction of a second off every copy of the graphic can add up to quite a lot of time overall.

© Global Graphics 2021. All rights reserved.

12.7 Use standards when they're relevant

Over the last couple of decades a number of standards have been developed in the International Organization for Standardization (ISO) and elsewhere that can assist in making variable data print workflows more efficient.

- PDF/X helps to enforce best practice, especially in color and font usage. Creators making PDF for print in any sector with ink sets based on CMYK are recommended to make those files PDF/X compliant.

- PDF/VT is built on top of PDF/X, adding extra requirements and recommendations specific to variable data printing. We recommend that you make use of the 'hints' defined in the PDF/VT standards when you have the information necessary to do so; in some cases that will require a bit of work by the composition tool vendor to support them. See also *Appendix A – Hints to enhance performance in Harlequin RIPs,* which describes specifically how Harlequin RIPs make use of those hints.

- If your job includes technical marks such as cut lines, fold lines, dimensions etc. we recommend that you identify them as such using the PDF Processing Steps standard (ISO 19593-1). It's a recent standard, and so is not all that widely implemented yet, but it's going to be hugely useful in optimizing workflows and avoiding errors in the DFE, so it's worth getting ahead of the game with it. But, like PDF/VT hints, in most cases the composition tool vendors may need to do some work before an operator can tag marks as processing steps.

It's worth remembering that a PDF/VT file must be a fully valid PDF file, and a fully valid PDF/X file as well, and that adding PDF Processing Steps also doesn't stop a file being a fully valid PDF. So even if you've just been asked to send an optimized PDF file, sending a PDF/X or PDF/VT file is likely to be useful by helping you reinforce your own self-discipline in managing colors and fonts.

You may also find the "PDF/X-plus" specifications from the Ghent PDF workgroup (gwg.org) useful in avoiding problems with unprintable graphics in your files.

There are some notes on several helpful standards set out in *Appendix B – Relevant standards.*

© Global Graphics 2021. All rights reserved.

12.7.1 Looking forward

At the end of 2020 three new standards were published: a „dated revision" of PDF 2.0 (think of it as a 2nd edition with some minor corrections to the text and a few new small features), a new PDF/X-6 standard, based on PDF 2.0, and a new PDF/VT-3 standard, based on PDF/X-6. These are intended, and expected, to become adopted as the mainstream formats for preparing files for production printing. But it will take a few years to achieve that, so some care should be taken through 2021 to ensure that you're not sending files that use a later standard than the printing company or converter can handle.

12.8 Optimizations in VDP workflow software

The recommendations above are relevant for graphic designers and composition operators in at least some workflows. But there are some optimizations that can only be addressed by the software vendors involved, either in asset creation and management, or in the composition tools themselves. These tend to be deeper into the technical aspects of exactly how a PDF file is constructed.

A composition tool vendor should also review the recommendations above, because they may well have an opportunity to help operators achieve the desired goal, and therefore to make the output from their tool more efficient at the printing stage. That may, for instance, be by offering the ability to discard clipped regions of images, down-sample very high-resolution images etc.

12.8.1 Embed each image in the PDF just once

The data for each image is embedded in the PDF file as an Image XObject. The description of the graphical contents of each page then includes a pointer to the XObject to place that image on that page. If the same image is used many times within a single PDF file then the image data can be embedded many times … or it can be embedded just once and the pointer from the page descriptions can all point to that same copy.

If multiple copies of the same image are embedded in the PDF that will evidently bloat the file size. Less obviously it will reduce the efficiency of the VDP optimizations in some DFEs because the images may be seen as different and therefore each copy may be processed separately, increasing the work required unnecessarily and slowing the job down.

Whenever possible only one copy of each image should be embedded. If the same source image is used at multiple different sizes on the pages those may either use the

© Global Graphics 2021. All rights reserved.

same embedded copy or a separate copy at a suitable resolution may be used for each final size.

12.8.2 Don't tile or stripe (most) images

A couple of decades ago it was common to write images into page description languages as a series of rectangular tiles, or as strips.

DFEs and RIPs at that time didn't have access to much RAM, and the intention was to ensure that the RIP didn't need to hold very large amounts of image data at the same time. RAM costs are still a factor in DFE design, but the amounts now used are many times higher than they were back then, so this 'workround' is no longer required.

On the other hand, there is a measurable cost for the RIP to set up and tear down a processing pipeline for each image, so making the DFE handle a large number of small images instead of a single large one can make it run slower.

One extreme example of inefficient practice can often be seen when an image has been placed on a page in a design application and then a single color in the image has been marked as transparent by the user. Some applications will generate a huge number of very small images, often in strips only one pixel tall, in the page description language. If they were to include the whole image as one, and to use a stencil mask or color key mask on that image it would increase processing speed in the DFE hugely.

And that slow-down is sometimes multiplied by encoding the image strips as in-line images instead of image XObjects. In-line images make it harder for the RIP to separate processing images from that of the rest of the graphics within a page and therefore subvert some of the optimizations that might otherwise be applied.

The exception to this recommendation is when processing very large images, either hundreds of megapixels, or covering very large physical areas. In those cases it can be more efficient to split into a few large images instead of one huge one.

12.8.3 Use a constant opacity rather than a soft mask with constant values

There are two ways of specifying how transparent a graphic should be within a PDF file: you can set a constant opacity value for fills and strokes (using the CA/ca keys), or you can attach a soft mask (SMask in the PDF, or within a JPEG2000 image). Soft masks can be very useful if the transparency should vary across the graphic, e.g. for softening the edges of an image. But we've also seen them used

© Global Graphics 2021. All rights reserved.

quite a lot where the transparency is uniform across the whole graphic. The most inefficient examples add a soft mask where all of the values are either 1.0 (indicating that the element is fully opaque) or 0.0 (indicating that the element is fully transparent and should not be visible at all).

If the element should be fully opaque the best way to represent that is to omit the SMask entry completely, or to set it to/None.

If the element should be fully transparent (not visible) then don't include it in the PDF file at all!

And if the element should have a constant transparency that is neither fully opaque, nor fully transparent, just use the CA or ca keys to set that value and omit the SMask key or set its value to /None.

12.8.4 Don't subset fonts

Some software subsets fonts when embedding them within a PDF file. It's a technique that was originally developed to reduce file sizes and to make it marginally harder to copy fonts by extracting them from PDF files. The incremental increase in file sizes to include a whole font in a VDP file is now trivial compared to disk sizes and communications speeds, with the possible exception of multi-byte fonts, for Japanese or Chinese for example. And most font vendors have adopted different models for font sales that don't rely on avoiding embedding them completely. So most of the advantages of subsetting fonts have disappeared.

On the other hand there are distinct costs from subsetting fonts in a VDP job if that is performed per page. Each subset of the font will be regarded by many RIPs as a different font. That means that the cache of rendered characters must be built from scratch for every different subset font, which slows the job processing down slightly.

So we recommend that you don't subset fonts in a VDP job or, if you do subset, you embed a single subset that includes all of the glyphs used on all pages for all instances (recipients, labels etc.).

If, however, you're generating personalized instances of a PDF file for web or mobile device delivery you may want to continue subsetting embedding fonts for each instance, especially if using multi-byte fonts.

© Global Graphics 2021. All rights reserved.

“ It’s difficult today to imagine a world in which business, commerce, retail and industry could function successfully without being able to utilize the extensive capabilities of variable text, variable graphics, sequential numbering and coding, track-and-trace, and the still, fast growing, fields of digital design, workflow and printing.

Yet all of this was still in its infancy just 50 years ago when I first became involved in research, writing, publishing reports and talking about bar coding, thermal printing, price-weight labels, and new innovations in print-on-demand using the early inkjet technologies.

Much has changed since then. The whole area of digital printing – using both toner and inkjet technologies – has evolved into the mainstream printing markets and new investment is coming close to dominating the world of label printing of the future.

All of this would not have been possible without a whole range of new high-tech companies emerging over the past 30 or 40 years to take the industry rapidly forward. Companies involved in digital printing technologies, design, graphics, pre-press, management information systems, workflow, file transfer, and much more, that both individually and together have so successfully transformed the VDP market.

Indeed, companies such as Global Graphics Software that has surely risen to become one of the world’s foremost experts in printing, PDF and digital document software, RIPs and print controllers. Their 2014 guide on optimal PDF file creation for variable data printing certainly enjoyed much success.

I’ve little doubt that this new, updated guide will also enjoy ongoing success in continuing to bring all those involved in VDP workflow ever closer together, and therefore ensure that the industry stays at the forefront of VDP innovation for many years to come. ”

Mike Fairley

© Global Graphics 2021. All rights reserved.

APPENDIX A

HINTS TO ENHANCE PERFORMANCE IN HARLEQUIN RIPS

Harlequin RIPs will benefit from well-made PDF and PDF/VT files, following most of the guidance in this book, but they can also act on some PDF/VT and product-specific hints for accelerating job processing. This appendix is, unlike the rest of this guide, specific to Harlequin RIPs.

The product-specific hints described here are defined for use in Harlequin but will not adversely affect processing in RIPs from other vendors. That means that there's no risk, and probably a benefit, in including them when you have the information to do so. When you make a file, you may not know whether the PDF file you're creating will be consumed by Harlequin or not. And even if the original plan was to use a different RIP, a late change in production planning, perhaps in response to a press problem or a previous job overrunning, may mean that it will be rendered in Harlequin after all.

With appropriate RIP configurations Harlequin will act on the hints described below; and it will do so whether the file is created as PDF/VT, or as 'baseline' PDF.

We recommend that you include the `GTS_Encapsulated` hint, defined in the PDF/VT standard, when you have the information available to do so.

PDF/VT 1 and PDF/VT 2 require that XObjects marked as `GTS_Encapsulated` must also be isolated transparency group objects if there is any live transparency anywhere in the file. Section 7.3 above recommends not using any live transparency if you don't have to, so marking something as a transparency group when it does not include live transparency might seem counter-productive. But Harlequin includes very efficient techniques to optimize a transparency group that doesn't actually use transparency, so there isn't any real cost to using the `GTS_Encapsulated` tag.

Harlequin uses the `GGSL_XID` entry in an XObject dictionary as an indication that the XObject should be treated as atomic, i.e. that it should not be split into multiple components.

© Global Graphics 2021. All rights reserved.

GGSL_XID is very similar to the GTS_XID hint defined in PDF/VT-1; the value of GGSL_XID should be a string representing a unique identifier (such as a GUID). Alternatively, you can simply set GGSL_XID to true to indicate that the XObject should be treated atomically; Harlequin will then calculate its own unique identifier for it. Using a unique identifier can be slightly more efficient if the same re-used graphics are used in multiple PDF files, e.g. in support of a pseudo streaming workflow where the job is sent in 'chunks'.

If there are forms in the PDF file that are re-used more than once and that should be used atomically, then we recommend that you set GGSL_XID in them. A Form or Image XObject used as the background for every copy of a piece of work would be an example where using GGSL_XID can help to increase the RIP's processing speed.

If a Form XObject is being used as a container for layout or imposition, and contains data that is only used once, possibly in conjunction with other data that is re-used multiple times, do not set GGSL_XID in it, because you won't want Harlequin to treat it as atomic.

GGSL_XID has been defined because the new PDF/VT-3 standard does not make use of the GTS_XID key. We also recommend that GGSL_XID is used instead of GTS_XID where you have the option to do so to avoid any possible clash in optimal use of GTS_XID in PDFs tuned for different RIPs. But if you cannot add GGSL_XID then use GTS_XID instead, as described in the PDF/VT-1 standard.

© Global Graphics 2021. All rights reserved.

FIG 34 shows a page, in dark blue, with four postcards imposed on it, in light blue. Each postcard would typically be represented as a Form XObject, just to make the imposition a little simpler when constructing the file. Each of those forms here would include a block of text and a reference to another form (shown as the white rectangle, with graphics that are re-used identically on every postcard). In this case the white form should include `GGSL_XID` to denote that it should be treated as atomic and not split up.

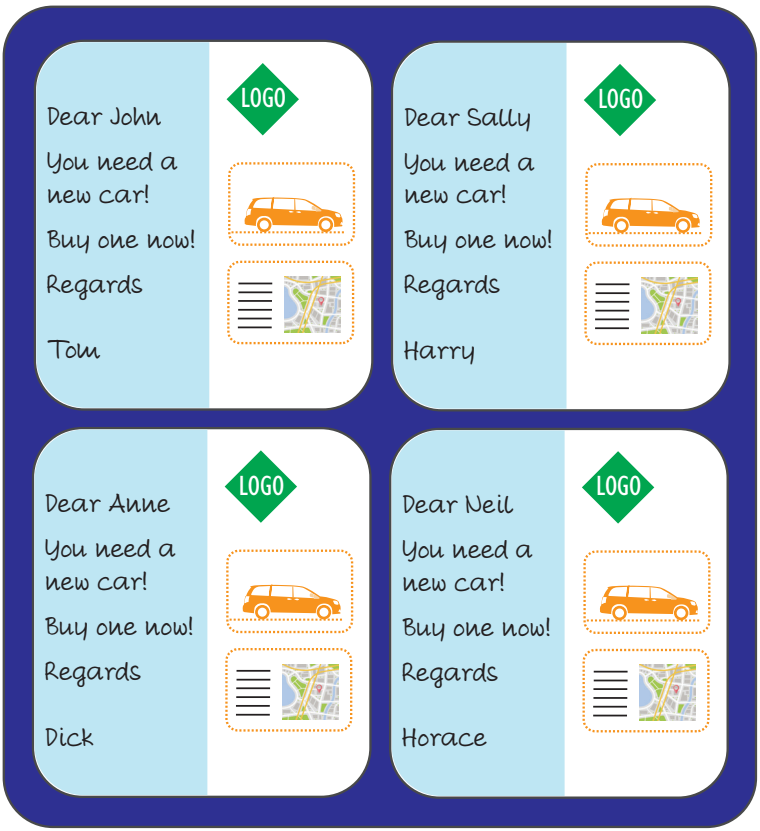


FIG 35. A simple page of imposed postcards.

© Global Graphics 2021. All rights reserved.

The four pale blue forms, used for each of the postcards, should not have GTS_XID or GGSL_XID set in them.

All five forms (both the postcards and the re-used graphics) should be constructed in such a way that it would be valid to tag it with GTS_Encapsulated=true and should be tagged as such.

Global Graphics has also defined the GGSL_Opaque hint, which takes a Boolean value (true or false). This should be set to true for all XObjects where it is known that:

- There is no live PDF transparency in the XObject; and
- There are no spot colors set to overprint in the XObject

Overprinting spot colors are listed here because, in most cases, they will be emulated using process colorants when printed on a digital press. The creator of a variable data PDF file may not always know exactly which spots will be emulated and which will be printed with real colorants, unless they're special spots, such as White, varnish, MICR, fluorescing, foiling, 'invisible' or similar. Harlequin's merging of emulated spots with other objects in the process color space uses the same process as transparency compositing so that emulation of overprints between that spot and other graphics can be accurately rendered.

Conversely, the GGSL_Opaque hint should be set to false if the XObject is known to contain live PDF transparency, or spot colors set to overprint.

If you are not confident that you have the information to set any of these hints according to the rules above, don't include the hints at all.

Harlequin supports additional hint keys for specific workflows defined by our OEM partners. If you need to achieve something unusual, please start by contacting your press vendor for advice.

© Global Graphics 2021. All rights reserved.

APPENDIX B

RELEVANT STANDARDS

Section 12.7 recommends using standards such as PDF/X and PDF/VT to support good practice in creating PDF files for variable data printing. This section provides some more information on relevant standards.

The PDF/VT standards are built on PDF/X which, in turn, are built on PDF. The Processing Steps and Print Product Metadata standards can be used in conjunction with several PDF/VT, PDF/X and PDF standards.

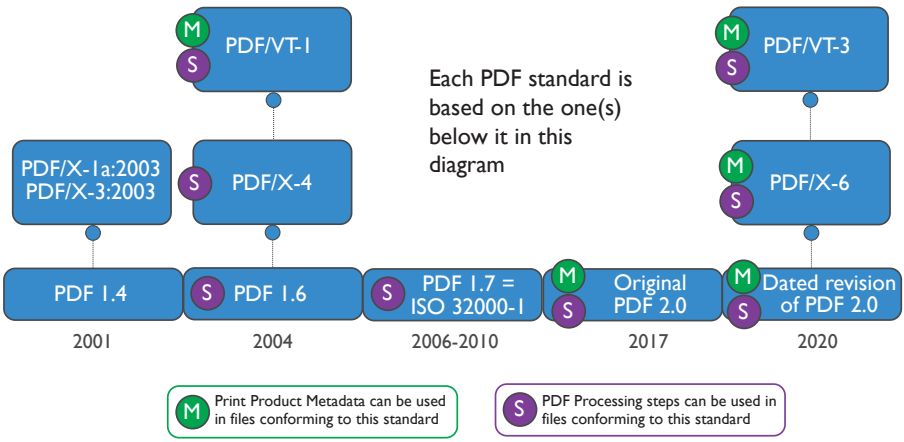


FIG 36: Relationship between print-related PDF standards.

14.1 PDF/X (ISO 15930)

PDF/X is a family of ISO standards; they are designed to encapsulate best practice in creating a PDF file that can be sent from one department or organization to another for production printing with minimum problems.

© Global Graphics 2021. All rights reserved.

The PDF/X standards mandate that certain items are included in a file when it's created to ensure best practice is followed. The various PDF/X standards all require that:

1) All fonts required in the document are embedded, avoiding problems with missing fonts or the use of a different version of a font by the same name at the print service provider.

2) The colors used for all objects must be defined sufficiently completely that they can be reproduced consistently and accurately.

3) The color reproduction of the output device for which the job was designed is specified, allowing accurate and consistent proofing and emulation of the job on other devices, and preflight to identify upstream mistakes quickly and easily.

The real value of PDF/X standards, however, isn't obvious from a simple list of technical prohibitions, requirements and recommendations. It boils down to two things:

a) Application vendors can add an option to export to a specific PDF/X conformance level, and users can then create files that are highly likely to print well on a press run by some other company, without knowing exactly what that conformance level requires. Even technically savvy users can benefit because that same option will reinforce self-discipline in following best practice.

b) A print service provider, converter, or other company printing for other people can say very simply "give me a PDF/X-4 file" and know that they will almost certainly be able to process what they get sent. The designer, upstream prepress operator or composition tool operator can translate that request into action because of point a) above.

A PDF/X file also fully conforms to the PDF specification, so if you're creating files you can make them as PDF/X in order to gain the benefits of automatic checking of things like font embedding, even if you've just been asked for a PDF file.

© Global Graphics 2021. All rights reserved.

14.1.1 PDF/X conformance levels

The PDF specification itself has evolved over the last couple of decades, as has the demand for richer graphics in printed work, and for formats to support increased automation in the print workflow. As a result there are now a number of PDF/X 'conformance levels' to choose from:

PDF/X-1a – first published in 2001, and then re-issued in 2003, PDF/X-1a is now regarded as a very conservative format. It's based on PDF-1.4 and all color data must be expressed in CMYK or spots; no device independent color spaces (ICC), RGB or Lab data are allowed. It also prohibits live PDF transparency and doesn't allow constructs added in later PDF specifications, such as PDF layers (optional content).

PDF/X-2 – didn't take off and is best avoided.

PDF/X-3 – first published in 2002, and then re-issued in 2003, PDF/X-3 is a slightly more open standard than PDF/X-1a. It's still based on PDF-1.4 and prohibits transparency and PDF layers, but it does allow device independent color spaces.

PDF/X-4 – first published in 2008, and then re-issued in 2010, PDF/X-4 is based on PDF-1.6 and allows device independent color spaces like PDF/X-3. It also allows live PDF transparency and PDF layers to be used, with some appropriate limitations. At the time of writing PDF/X-4 is the most obvious format to use for delivering the majority of PDF files for production printing.

PDF/X-4p and PDF/X-5 – are specialist variants designed for specific workflows where the receiver of the file has access to additional data that must also be available to process the files correctly. These standards should only be used if the company who will be receiving the PDF files asks for them.

PDF/X-6 – Is a new PDF/X standard published in November 2020 and building on PDF 2.0 (see 14.3). It takes advantage of new functionality in PDF 2.0 such as page-level output intents, and also adds value in multi-channel delivery because of the extended accessibility support in PDF 2.0. In time this will develop into the obvious choice, but use care through 2021 because not all printing companies or converters

© Global Graphics 2021. All rights reserved.

will necessarily have upgraded to support PDF 2.0 across all of their workflows yet.

PDF/X-6p and PDF/X-6n – like PDF/X-4p and PDF/X-5, these are specialist variants designed for specific workflows and should only be used if the company who will be receiving the PDF files asks for them.

14.2 PDF/VT (ISO 16612-2)

In 2010 the International Organization for Standardization (ISO) published a new standard called "ISO 16612-2:2010 – Graphic technology – Variable data exchange – Part 2: Using PDF/X-4 and PDF/X-5 (PDF/VT-1 and PDF/VT-2)". It's designed specifically to support robust delivery and production of modern variable data print jobs; the 'VT' in the name standards for "Variable and Transactional". A new standard, ISO 16612-3, has just been published at the end of 2020.

By building on PDF/X, and therefore on PDF, these standards enable the use of many of the features that graphic designers have come to expect to be able to use for work in commercial print, publication, etc., and therefore wished to use for complementary advertising in direct mail and transpromo campaigns, and in labels and packaging. By also including document metadata that can convey the designer/purchaser's requirements, it allows far more complete automation of production in support of today's increasingly complex and demanding requirements around page count and separate components to be delivered together.

14.2.1 PDF/VT conformance levels

Between them, the two PDF/VT standards define four conformance levels, reflecting both different use cases for variable data, and changes in the print industry over time.

PDF/VT-1 – all content for a print job is included in a single PDF file, which must also conform to PDF/X-4 (ISO 15930-7:2010). The vast majority of current PDF/VT production is PDF/VT-1, and until the publication of PDF/VT-3 at the end of 2020, this was the only PDF/VT standard that we would recommend for current workflows unless all parties in that workflow agree to use one of the others.

PDF/VT-2 – designed to support a 'chunking' workflow to allow something almost indistinguishable from streaming, i.e. where the first pages of the job are being

© Global Graphics 2021. All rights reserved.

printed before the last ones have been created by the composition engine. It does this by providing a method whereby large assets such as images that are used multiple times (e.g. for many recipients each) can be saved into a single PDF file, known as a target file. A series of 'chunks', each defining a range of pages to be printed and saved as a PDF/VT-2 file, is then produced. Each PDF/VT-2 file includes references to the assets in the target file(s), which means that those large assets don't need to be repeated in every PDF/VT-2 file. PDF/VT-2 is not widely implemented or used.

PDF/VT-2s – is a variant of PDF/VT-2 where both the target files containing re-used assets and the PDF/VT-2 files themselves are wrapped into a single MIME stream. The intention is to simplify delivery of a stream for printing where there isn't a shared file system accessible to both the submission tool and the DFE. PDF/VT-2s is even less widely implemented than PDF/VT-2 and should be avoided.

PDF/VT-3 – Was published in late 2020 and is based on PDF/X-6, which, in turn, is based on PDF 2.0. Amongst other things this allows better color management of jobs that are printed on multiple different media, and allows a unification of PDF generation for multi-channel delivery with excellent accessibility capabilities; see 14.3. Just like PDF/X-6, PDF/VT-3 is expected to become the most commonly used PDF/VT standard over time, replacing PDF/VT-1. But care should be taken throughout 2021 because not all press DFEs will necessarily have been upgraded to support PDF 2.0 yet.

Just like PDF/X, the real value of PDF/VT is more in simplifying communication of requirements and best practice than in defining anything significantly different from what can be achieved in baseline PDF. In a sense it relieves the graphic designer and composition tool operator of the need to consider some of these constraints when they make a file; just select "PDF/VT" in the menu when generating the file for print and it will be done for you.

But the PDF/VT standard concentrates on providing support for predictable and repeatable output and for automation; it does not focus on how the desired elements should be written into that file in order to maximize the efficiency of processing.

© Global Graphics 2021. All rights reserved.

So using PDF/VT is a very good way of improving the document delivery workflow in many ways and is definitely recommended. But it's not the whole story. There are many things that users can do to optimize processing of those jobs as well, and to help avoid last-minute problems. Those are the subject of this guide, and most are equally applicable to both PDF/VT and 'baseline' PDF.

14.2.2 Key advantages of PDF/VT

Using PDF/VT files instead of pragmatically defined "optimized PDF" files provides a number of distinct benefits for both creators and printers:

a) PDF/VT builds on the work done for static artwork delivery for both conventional and digital print in the PDF/X family of standards, which have become an extremely common way of enforcing best practice and simplifying the creation of pre-flight profiles etc.

b) PDF/VT provides the framework for a composition engine to include a hierarchical tree of metadata in the file, to encapsulate the intents and expectations of the designer/purchaser. *See 14.4* for more detail.

c) A PDF/VT file may include hints that can be used in the RIPs within a DFE to assist in optimizing VDP processing. *See Appendix A* for notes on how Global Graphics' Harlequin RIP processes hints in PDF/VT and in optimized PDF files; RIPs from other vendors may have equivalent support..

14.3 PDF 2.0 (ISO 32000-2)

ISO took over maintenance and development of the PDF specification from Adobe in 2008, and the first standard version of PDF was issued as ISO 32000-1 in 2010. That standard is nominally (but not exactly) the same as PDF 1.7, differing mainly in being much clearer in distinguishing recommendations from requirements. In 2017 the first update fully developed in the standards community was published, as ISO 32000-2. Files conforming to the new standard are referred to as PDF 2.0.

The new standard includes many clarifications and improvements to the text. It's also been expanded and improved significantly in areas such as encryption and

© Global Graphics 2021. All rights reserved.

accessibility. But for the purposes of this guide the most important change is that it now allows output intents to be defined for every individual page instead of only once, at the document level.

Output intents were first defined for use in PDF/X standards and carry a statement of the characterized printing condition which the designer assumed would be used when the file is printed. It does this by name, e.g. using the names of the characterizations defined on the ICC's web site (color.org), and often by inclusion of an ICC color profile, which should be used in many cases when processing the file for print.

When PDF/X was first created it was anticipated that the most common use case would be in supplying display advertising for magazines and newsprint, and therefore that files would only contain a single page. So only one output intent was specified, and it applied to the whole PDF file.

PDF/X has obviously been adopted much more widely than that, and some variable data print cases are difficult to deliver using that original model. The most obvious challenge is variable data print jobs using multiple different substrates, because a different output intent should be used for each substrate.

PDF 2.0 has adopted and extended the original PDF/X output intent, and it's now possible to include an output intent on every page. So now one output intent can be created for each substrate, and those can be referenced from each page so that the right color management is triggered.

PDF 2.0 has also adopted the DPart structure used for print metadata from the first PDF/VT standard (see 14.2 and 14.4), and a considerable amount of work has been done on areas outside of production printing workflows, such as better support for accessibility tools such as screen readers. If you're doing any form of multi- or omni-channel publishing you may find that additional accessibility support very useful, e.g. to achieve Section 508 compliance, and the equivalents in other countries.

At the end of 2020 a new "Dated Revision" of the ISO PDF 2.0 standard was published. The primary aim of this new publication was to clarify additional areas

© Global Graphics 2021. All rights reserved.

and to resolve problems that had been encountered in real-world usage, not just of the 2017 PDF 2.0 standard, but also of previous PDF specifications.

Both PDF/X-6 and PDF/VT-3 are based on this new dated revision, and you should use this latest document rather than the 2017 edition whenever possible.

So PDF 2.0 adds value to many workflows, including those for production printing, but it does also bring a small amount of risk. If a file has used some of the new features in PDF 2.0 those will usually be silently ignored by an older reader. PDF was designed to be very flexible, and to allow custom and proprietary data to be embedded virtually anywhere in the file structure. It does that by saying that a reader should simply ignore anything it doesn't recognize. To a PDF 1.7 reader, most new PDF 2.0 features are just objects that it won't recognize and should therefore ignore. The most common exception to that rule is around security; if the PDF file uses the new AES-256 security introduced in PDF 2.0 then an older reader will probably be unable to read that file at all.

The biggest risk area for people considering when and how to roll out PDF 2.0 support is therefore that a PDF 2.0 file using new features may have those new features silently ignored. Some readers will emit a warning that the file you're opening has a PDF version number that is not explicitly supported. That's helpful, but it can never be more than a hint to take care because that older reader doesn't know anything about any new features in the file; it cannot possibly know if they're important to you or to your workflow.

This means that the safest approach to adoption of PDF 2.0 is to ensure that all applications and tools that consume PDF are upgraded to support PDF 2.0 before you start thinking about upgrading the file creators. Start at the end of the workflow and work upstream.

For more detail on PDF 2,.0 in production printing workflows, take a look at https://www.globalgraphics.com/download_file/33/217.

© Global Graphics 2021. All rights reserved.

14.4 PRINT PRODUCT METADATA FOR PDF (ISO 21812-1)

PDF/VT provides a framework for a composition engine to include metadata relating to pages and groups of pages in a hierarchical tree in the file. This data can be encoded into a structure known as the "document part metadata" or DPM.

The value of the DPM varies significantly between different print spaces and is likely to be most useful in print on paper, especially in the transactional and direct mail market. In those contexts it may, for instance, include information on the state and ZIP code of every recipient to allow post-composition selection or sorting.

It can also include details of how pages in the PDF file relate to complex deliveries to recipients, such as a combination of addressed envelope, cover letter and personalized catalog.

This means that the various components can be split across presses, or that the print service provider's workflow could be automatically configured on the fly, e.g. by interaction with a job ticket defined in the DFE itself. This, in turn, can be used to control how the job is imposed, printed and finished.

For simple VDP jobs it's relatively easy to configure the press and finishing line manually as required, but that becomes more challenging as jobs get more complex. That complexity might be because there are multiple components for every recipient (e.g. a cover letter and personalized booklet using one stock for the cover and another for the book block), or even just a different number of pages for each recipient, especially if those pages need imposing for printing. Even if you're only printing simple jobs, if they are short and don't all need exactly the same equipment configuration, adjusting for each job can become a significant time-sink and source of mistakes.

The DPM allows all of the information required to process a job to be encapsulated within the job itself, enabling effective automation in the print-shop. In development of PDF 2.0 the ISO PDF committees recognized the value of DPM and adopted it from PDF/VT into PDF 2.0 itself. But the PDF/VT and PDF 2.0 standards only define the DPM structure that can be populated; they don't define a specific language to

© Global Graphics 2021. All rights reserved.

use within that structure. That is where the "Print product metadata for PDF files" standard comes in. Published as ISO 21812-1 in 2019, it builds on the concepts developed over the last decade or so as 'intent' elements in JDF, although it does not use the JDF XML structure to do so. Intent elements describe what the final printed and finished piece should look like, but don't specify the processing that should be done to achieve that result. It's therefore vendor and workflow neutral, in that it allows the prepress or press operator, or the software in the workflow itself, to decide how to create the desired output.

Adopting the JDF concepts makes use of the experience developed in real-world systems over the past couple of decades, and also means that values from the DPM structure in a file can be very easily mapped into a job ticket if the DFE or workflow consuming the PDF file happens to use JDF for that job ticket. Even if the DFE is not using JDF, it still means that the information carried by the DPM is clear and well defined.

Combining print product metadata with an optimized PDF or PDF/VT file means that the creator (or some other component in the workflow between creation and processing for print) can identify the media to be used, how the job should be folded and bound, along with many other requirements.

To come back to an example that can be slightly challenging in current systems, it's relatively easy to tag page ranges in a PDF file for a job that includes, for instance, a

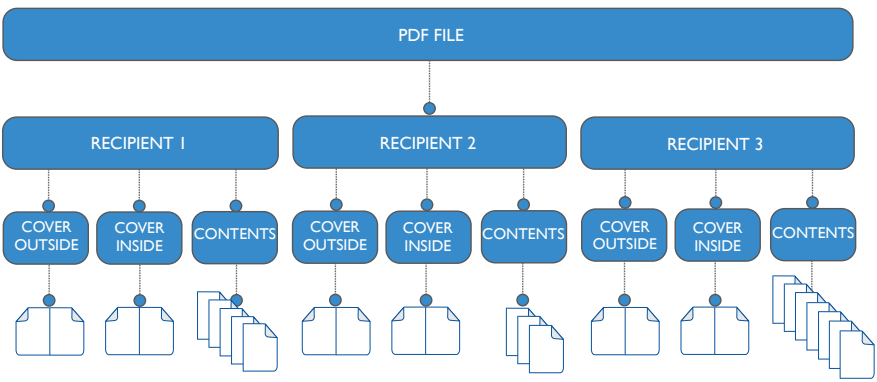


FIG 37: Hierarchy of pages in a personalized saddle-stitched booklet.

© Global Graphics 2021. All rights reserved.

saddle-stitched booklet for each recipient, even if the cover of the book should be printed on a different stock and each booklet contains a different number of pages.

The print product metadata must obviously be consumed by the print workflow, which may mean inside the DFE for a digital press, or may be slightly further upstream, to enable a job to be split over multiple presses, and for managing off-line finishing as well. As a new standard, systems to do that are not yet widespread.

When it is consumed, the print product metadata may be simply presented to an operator so that they are aware of how the designer/purchaser expected the job to be printed. In a more sophisticated workflow the metadata may be merged with a pre-existing job ticket, either fully automatically, or by offering the operator additional information in support of their configuration of the system.

14.5 PDF PROCESSING STEPS (ISO 19593-1)

Historically several methods have been used to identify technical marks such as cut and fold lines, dimensions etc., in a PDF file. Sometimes that's by using a spot color, perhaps called CutContour, or by encapsulating them into a PDF layer (optional content group) with some suitable name.

But both of those methods have one significant drawback, in that the name used for the spot color or for the layer is not standardized. Operators can, and do, use whatever name is common practice in their company, and in their language.

When the company who will print the job receives the PDF file, they need to look inside it to work out what those names are so that they can configure the DFE appropriately. For a proof some people will print all of the marks, but for the production print the cut lines, and anything else that might fall inside the live area of the job, must be turned off to avoid spoilage.

That inspection and configuration takes time and creates an opportunity for making mistakes and even proprietary metadata in PDF files often only identifies spot colors as "technical", without providing sufficient detail for automatic processing.

© Global Graphics 2021. All rights reserved.

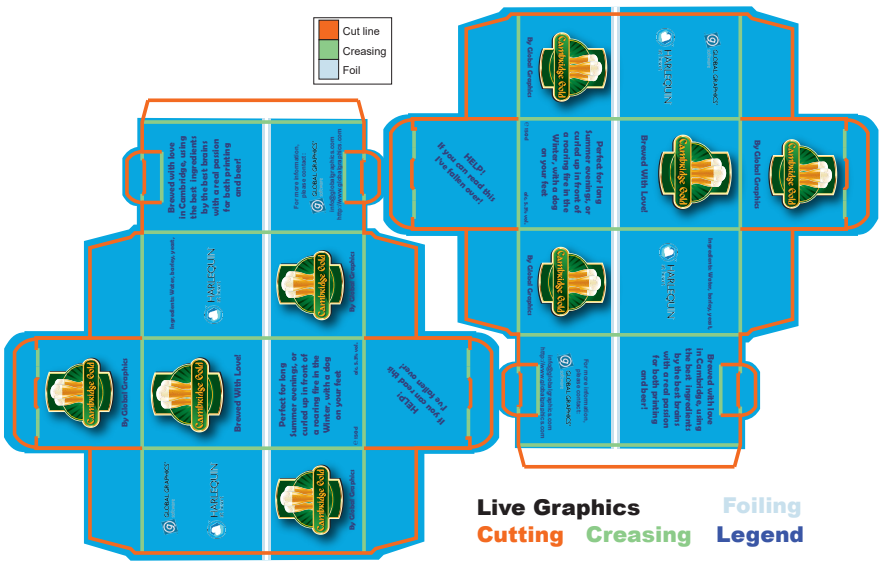


FIG 38: Common technical marks for a premium folding carton.

The PDF Processing Steps standard, published in 2018, builds on the idea of using PDF layers, but adds a standardized metadata structure to tag those layers as specific forms of technical marks. So it no longer matters whether the creation operator called a spot color or layer 'Cut', or 'Die' or 'Die line' or 'Dieline', or 'Coup' etc.; the consuming software can look at the metadata and know that it's a cut line. And that, in turn, means that a DFE on a digital press can be configured to turn off cut and fold lines etc., knowing that any file conforming to PDF Processing Steps will be processed as desired, with no review of spot color or layer names required.

It's a very new standard, but it's already implemented in software from Callas, Esko, Global Graphics and HYBRID Software.

© Global Graphics 2021. All rights reserved.

GLOSSARY

AFP/IPDS	A file and data stream format used primarily in transactional printing, and for very large direct mailing jobs for the banking and financial industry. See 7.
Asset	One graphical element used in constructing a variable data printing job.
bpp	Bits per pixel, describing the 'bit depth' of a **raster.** The output from a RIP may be 1bpp if it represents only no ink or 100% ink; it may be 8pp or 16bpp if it contains continuous tone (unscreened) data. For presses using multi-level **screening** it may also be 2bpp or 4bpp.
Brand colors	Colors that a brand regards as important for recognition by customers; typically includes colors used in logos and the principle colors used in product packaging designs. Usually represented as spot colors in a PDF file.
CMYK	The set of ink colors used most often for printing: Cyan, Magenta, Yellow and Black. Also used for a family of subtractive color spaces using those same colorants. A color space for a specific press/ substrate combination, or an exchange color space such as Fogra 42L is a member of the CMYK color space family.
CMYKOG, CMYKOV, CMYKOGV	Extended gamut ink color sets, adding two or three of Orange, Green and Violet to **CMYK** in order to be able to accurately print a larger proportion of brand colors.
Clipping path	A shape used to hide parts of another graphic (most commonly an image) that fall outside of that shape.
Coding & Marking	Reliable and high-speed printing of very simple graphics such as serial numbers, sell-by dates etc., usually tightly integrated with a finishing, converting or filling line rather than with a printing production line, and with relatively low quality requirements. See 5.
Composition tool	A software application that combines graphical designs with data and rules to generate a file or stream for a variable data print job. See 9.
Controller	The server on which the software directly associated with a digital press runs; this terminology is most commonly used in transactional and direct mail as opposed to **Digital Front End.**

© Global Graphics 2021. All rights reserved.

DFE	See ***Digital Front End.***
Digital Front End	The server on which the software directly associated with a digital press runs; this terminology is used in most print sectors.
Direct Mail	Marketing collateral personalized for individual recipients and delivered through the postal system.
dpi	Dots per inch; a metric commonly used for the resolution of a digital press.
Drop shadow	A graduated tone added on one or more sides of a graphical element in a design to give an impression of depth. Drop shadows usually make use of live PDF transparency. *See 12.2.*
ECI RGB	An **RGB** color space with a wide gamut, meaning that it can represent very saturated colors as well as neutrals. Defined by the European Color Initiative (ECI).
EPID	"Every Page is Different"; see ***Every Page Different*** below
Embellishment engine	A machine designed to embellish printed output by adding varnish, tactile effects, foiling etc. *See 6.*
Engine speed	The speed at which a printing press can run when there is no limitation on the delivery of data to it.
Every Page Different	A class of jobs in which every page is different, and where those differences are not amenable to the variable data optimization methods described in this guide because there are no graphics shared between pages. *See 5.6.*
Extended Gamut	The color gamut of a printer or press describes the range of colors that it is physically capable of reproducing. An extended gamut ink set or press has a larger gamut, and can therefore print some colors that could not otherwise be printed, which can be important for accurately reproducing brand colors. *See 6.*
Flattening	Recreating the graphics in a PDF file in such a way that all use of live transparency is removed while (at least nominally) retaining the same visual appearance. *See 12.2.1.*

© Global Graphics 2021. All rights reserved.

Halftone	See **Screening.**
Hybrid press	A printing press that incorporates stations using a variety of different printing technologies. See *10.*
Image	A graphic represented as a grid of pixels, most commonly created by a digital camera or scanner. See *12.1.*
Imprinting	Mounting a digital printing unit on a printing press, finishing line or converting line to add variable data. See *10.*
Industrial print	In this guide "industrial print" is used to describe printing where the print forms part of the product, as opposed to delivering information or being part of the packaging for a product. It includes printing flooring, wall covering, laminates, textiles, ceramics etc. This term is, however, frequently used with other meanings, most commonly all production printing (rather than office or home printing).
Lab	One of a family of device-independent color spaces that can be used to maximize the accuracy of color reproduction, regardless of the press and substrate used for printing.
Layers	The term commonly used for optional content within PDF files.
lpi	Lines per inch; a metric used to describe how fine some types of **halftone screen** are on a print. Most screens used on digital presses should not be described in terms of lpi.
Mask	A raster image used to hide parts of some other graphic; very similar to masks as used in Photoshop.
Metadata	"information about information". In the context of this document that means data within a PDF file that provides information about how it was created and how it should be processed for printing and finishing.
Overprinting	Instructing a PDF RIP not to knock a graphic out of anything previously painted in the same position on the page if it is drawn in different colors.
Optional Content	The technical term for Layers within a PDF file. Optional content can be used to control the visibility of graphical elements on a page, commonly in production printing for versioning, or for technical marks.

© Global Graphics 2021. All rights reserved.

PDF	Portable Document Format; the file format currently defined by ISO standard 32000.
PDF/VT	A "PDF subset standard" describing constraints on PDF for delivering files for variable and transactional printing. *See 14.2.*
PDF/X	A "PDF subset standard" describing constraints on PDF for delivering files for production printing. See *14.1.*
ppi	Pixels per inch; a metric to describe the resolution of an image. In the context of this guide that's the effective resolution of the image at its final size on a page.
ppm	Pages per minute; a metric to describe the speed of a printer or press. It's most commonly used for sheet-fed digital presses used for print on paper; other sectors use feet or meters per minute, square meters per minute (or hour) or sheets per hour.
PURL	Personalized URL; an internet link created for a specific recipient of a specific variable data job to allow that recipient to access relevant data and a marketing department (or other job commissioning group) to identify when the recipient views that link.
Personalization	Generation of content that is designed for one specific person; examples include direct mail or labels printed for gift packs that have been ordered online. *See 5.4.*
Processing Steps	Standardized identification of the purpose of graphics in a PDF file that are not primarily intended to be printed. See *14.5.*
RGB	A family of additive color spaces using three colorants Red, Green and Blue. Typically used for photographic images. **sRGB** and **ECI RGB** are specific color spaces within the RGB family.
RIP	Raster Image Processor; an application that translates from a page description language such as PDF into a form that can be sent to inkjet heads or to the drum on a toner-based printer. Those don't understand fonts or even stroked paths and can only accept a **raster** in the right format.

© Global Graphics 2021. All rights reserved.

Raster	A representation of a graphical design as a rectangular grid of pixels. Images from a scanner or digital camera are rasters, but the term is used in this guide mainly for the output of a RIP, to be delivered to the inkjet heads or toner drum. Rasters are characterized by what colorants they include, the number of pixels across and down, the resolution (in **dpi**) across and down and the bit depth (in bpp).
Rendering	The process of generating a raster from a page description language such as PDF. Originally the primary purpose of a **RIP**, although modern RIPs include many more functions.
Screening	Very few printing devices can print a continuous ranges of amounts of ink in different locations on the substrate, so graphics must be simplified to reduce the number on ink levels to be applied. This is done using screening techniques that fool the eye into seeing all of the different colors that the design required and minimize the visibility of the screening pattern applied. On an offset or flexo press that means reducing to only no ink or full ink in any one location (for each ink color); many digital presses can apply a small number of different levels of colorant to each location, and therefore use 'multi-level screening'.
Shells	Pre-printed copies of a background design, over which a digital press can add variable data to generate final prints that incorporate both parts.
Soft mask	A **mask** which is not fully opaque, allowing some amount of any graphics behind the masked design to show through. Sometimes used to soften the edge of an image to avoid a hard edge.
sRGB	An **RGB** color space defined to represent a lowest common denominator of computer displays, which means it has a rather small gamut and cannot represent very saturated colors in the same way as, say, ECI RGB can.
Text	Words represented in a PDF file using a font. Note that a page may include objects that look like text, but where those 'characters' are saved as a filled **vector** outline, or as a pre-rendered **image**. RIPping speed (and the final appearance) can be affected quite significantly by which of these representations is selected.
Traceable	A printed piece is traceable if it includes sufficient detail to allow where and when it was printed, and by whom. *See* 5.2.

© Global Graphics 2021. All rights reserved.

Trackable	A printed piece is trackable if the brand owner, printer or other stake holders know where every instance has been delivered to. *See 5.3.*
Transactional	A personalized document representing something like a phone bill or credit card statement.
Vector graphics	Graphics in PDF files (like most modern page description languages) may be **images** or may be filled and/or stroked shapes and text. Of these, everything apart from images is described as vector graphics. Vectors can be viewed and rendered at any size or resolution and still retain smooth edges, unlike images which will normally start to show steps in diagonal edges as they are enlarged.
Versioning	The division of a large print job into several shorter runs with some variation between them. Labels or packaging for use across multiple geographical areas with different legal or language requirements is a common example. *See 5.1.*
White paper workflow	A workflow in which variable data printing is performed in a single step, rather than printing variable data on top of a pre-printed shell or imprinting. *See 10.*
Wide gamut	Another word for **extended gamut.** "Extended gamut" is most often used for the color gamut of a printing press, while "wide gamut" is more often used for the color space used in an **image**, e.g. for **ECI RGB.**
XObject	A structure used in PDF files to encode graphical elements that should be kept separate for any reason. That's often so that they can be written into the PDF file once and then used many times, saving space and increasing efficiency.

© Global Graphics 2021. All rights reserved.

About Martin Bailey

Martin first joined what has now become Global Graphics Software in the early nineties, and has worked in customer support, development and product management for the Harlequin RIP and is now a Distinguished Technologist. During that time he's also been actively involved in a number of print-related standards activities, including chairing CIP4, CGATS and the ISO PDF/X committee.

He's currently the primary UK expert to the ISO committees maintaining and developing PDF, PDF/X and PDF/VT.

© Global Graphics 2021. All rights reserved.

Impact Through Design with HP SmartStream Designer

HP SmartStream Designer is a plug-in for Adobe Illustrator® and Adobe InDesign® that allows you to create sophisticated personalized and mass-customization jobs which are printed on HP Indigo digital presses.

The plug-in includes exclusive features like HP Collage and HP Mosaic that let you create unlimited design variations and deliver brand differentiation and a high business impact.

To download and for more information, please scan the code or go to:
hp.com/go/smartstreamdesigner

Our mission is to provide cogent news and analysis about trends, technologies, operations, and events in all the markets that comprise today's printing and sign industries including commercial, in-plant, mailing, finishing, sign, display, textile, industrial, finishing, labels, packaging, marketing technology, software and workflow. Join here:

https://whattheythink.com/join

WhatTheyThink

PrintingNews Wide-Format&Signage

EMBRACE THE FUTURE OF CONNECTED PRINT AND PACKAGING

Take your business to the next level with interactive print. Deliver personalized experiences directly to consumers.

digimarc.com/industry/premedia

DIGIMARC DISCOVER®
(dmrc.app)

SCAN AD TO LEARN MORE

DIGIMARC

elan series

500 Images per minute

18" x 26" Maximum Sheet Size

1600 x 1375 DPI CMYK

Unrivaled Uptime, Unparalleled Support

Low Total Cost of Ownership

Driven by Global Graphics Harlequin Technology

Colour & Monchrome Inkjet Systems
Custom Solutions Development

DELPHAX SOLUTIONS
Superior Digital Printing

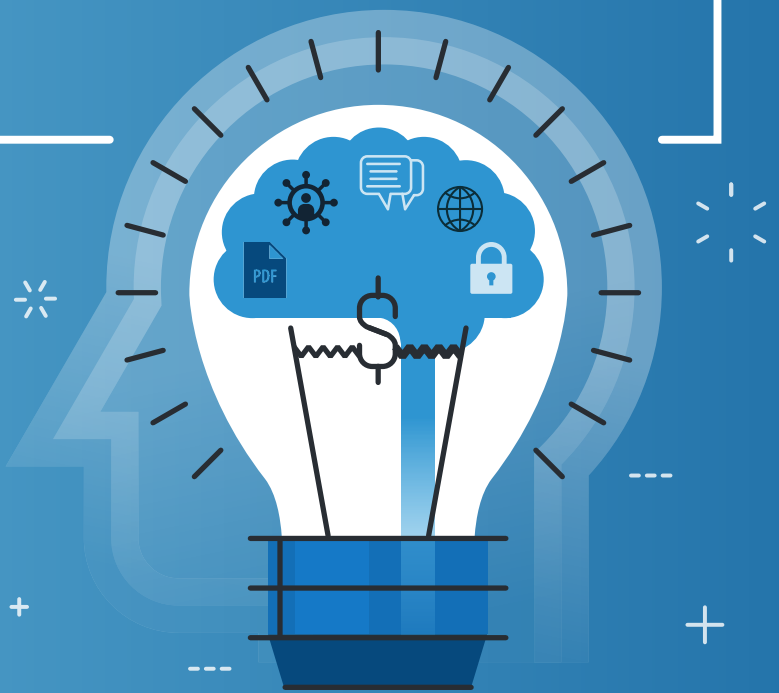
2810 Argentia Road, Unit 6
Mississauga, ON, L5N 8L2, Canada
+1-833-DELPHAX | www.delphaxsolutions.com | sales@delphaxsolutions.com

racami™

(678) 730-7700 | sales@racami.com | www.racami.com

- Securely Store PDF and Data for e-Delivery
- Facilitate Online Payments
- Manage Customer Preferences
- Optimize PDF for Printing
- Improve Document Production & Distribution
- Give Employees a Single Source of Truth

IDEAS FOR **SMARTER COMMUNICATIONS** COME FROM RACAMI



THINK YOU KNOW INKJET? THINK AGAIN.

The new KODAK PROSPER ULTRA 520 Press, powered by Kodak's breakthrough ULTRASTREAM continuous inkjet technology, completely redefines inkjet production printing. This cutting-edge web press closes the gap with offset.

See for yourself. Request samples or a virtual demo:
kodak.com/go/prosperultra520press

©Kodak, 2020. Kodak, Prosper, Ultrastream and the Kodak Logo are trademarks of Kodak.

HYBRID
SOFTWARE

The smart way to integrate
PEOPLE, PROCESSES & TOOLS.

VDP
VARIABLE DATA PRINTING

The continuous growth of e-commerce and personalized label and packaging designs fuels the overall demand for digital variable data printing.

- Moves VDP where it belongs: from press room to press
- All-in-one solution for prepress and VDP – based on PDF
- Consistent process for any digital press
- Optimized PDF creation for on-the-fly printing at full speed
- Optional CLOUDFLOW module for automated industrial production

www.hybridsoftware.com

HP PAGEWIDE

WEB PRESS T250 HD WITH

HP BRILLIANT INK

THE FUTURE GETS REAL

Make a brilliant move

Introducing the HP PageWide Web Press T250 HD with HP Brilliant Ink, the new standard for high-volume production inkjet web presses. HP Brilliant Ink delivers a vivid color gamut with outstanding results across a broad range of substrates, including offset coated and uncoated media.

Enjoy the benefit of one press for all applications—commercial, direct mail, publishing, and transactional.

Find out more: hp.com/pagewidewebpressT250

© 2020 HP Development Company, L.P.

keep reinventing

Members of the Global Graphics group of companies

GLOBAL GRAPHICS software

Core technology for digital print

meteor

Electronics, software, tools and services for industry inkjet systems

XITRON

Intuitive pre-press solutions

HYBRID SOFTWARE

Innovative productivity tools for the graphic arts industry

GLOBAL GRAPHICS

www.globalgraphics.com

Distributed by

GLOBAL GRAPHICS

www.globalgraphics.com

Second edition January 2021

With thanks to our sponsors...

Indigo WhatTheyThink DIGIMARC DELPHAX SOLUTIONS racami KODAK HYBRID SOFTWARE PageWide Industrial