



Ghostscript and MuPDF Status

OpenPrinting summit April 2013

Michael Vrhel, Ph.D.
Artifex Software Inc.
San Rafael CA



Ghostscript overview

What is new and what is coming...

Color Architecture Details

MuPDF



The Basics

Ghostscript is a document conversion and rendering engine.

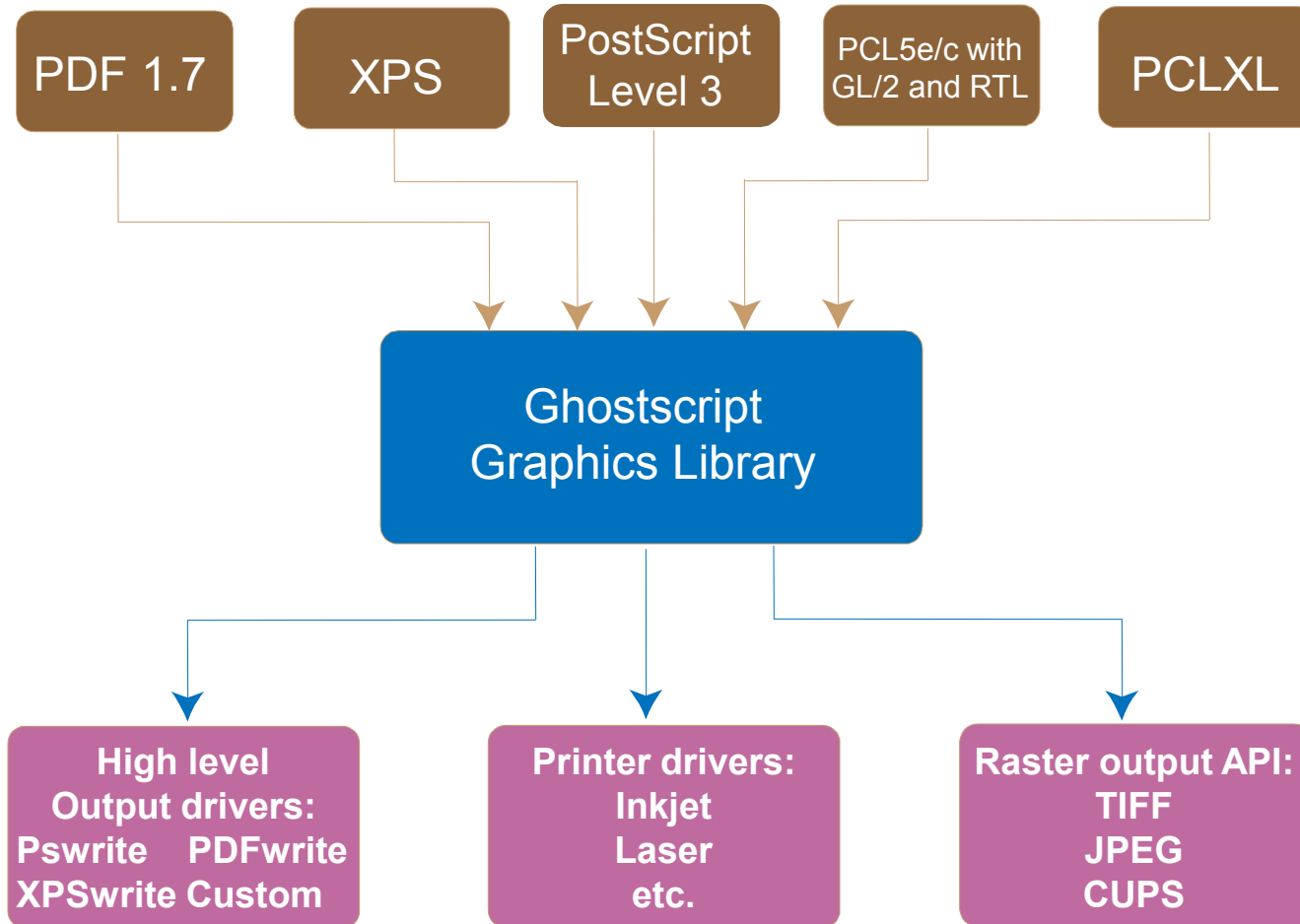
Written in C ANSI 1989 standard (ANS X3.159-1989)

Essential component of the Linux printing pipeline.

Dual AGPL/Proprietary licensed. Artifex owns the copyright.

Source and documentation available at www.ghostscript.com

Graphical Overview



Understanding devices is a major key to understanding ghostscript.

Devices can have high-level functionality. e.g. pdfwrite can handle text, images, patterns, shading, fills, strokes and transparency directly.

Devices may be set up to handle only certain high-level operations.

Graphics library has “default” operations. e.g. text turns into bitmaps, images decomposed into rectangles.

In embedded environments, calls into hardware can be made.

Raster devices require the graphics library to do all the rendering.

Relevant Changes to GS since last meeting....

Ghostscript can now use output intents defined in PDFs by using the "-dUsePDFX3Profile" command line option. (9.06)

tiffsep/tiffsep1/psdcmymk: Support for large numbers of separations improved. Removed reliance on the compressed color encoding. (9.06)

Ghostscript and GhostPDL distributed under the GNU Affero General Public License (AGPL). (9.07)

Ghostscript now has the option to be built as thread safe. Note that not all devices are thread safe. (9.07)

Relevant Changes to GS since last meeting....

The pdfwrite devices now supports linearized (or optimized for fast web view) output directly ("-dFastWebView") (9.07)

All interpreters now use Freetype by default to render all viable font types (9.07)

Ghostscript extended to support file sizes >4Gb - in particular reading and writing PDF files. (9.07)

Relevant Changes to GS since last meeting....

All CMYK devices can now support simulated overprint of spot colors using the "-dSimulateOverprint" command line option (9.07)

Support for use of DeviceN ICC color profiles as the output profile with the tiffsep and psdcmyk devices. (9.07)

Support for customized named color handling with DeviceN colors (9.07)

Support for black point compensation(9.07)



Relevant Changes to GS since last meeting....

Support for K preservation in CMYK to CMYK conversions. (9.07)

Support for DeviceLink profiles for graphic, image and text objects (9.07)

Support for custom color replacement (9.07)

Increased control in specifying color conversions as a function of object type (9.07)

LittleCMS updated to 2.4 (9.07)



Upcoming Changes to GS (release 9.08*)

Addition of option to interpret upcoming page on separate thread.

`-dBGPrint = true/false`

Detection of gray only content on page.

`-dGrayDetection=true/false`

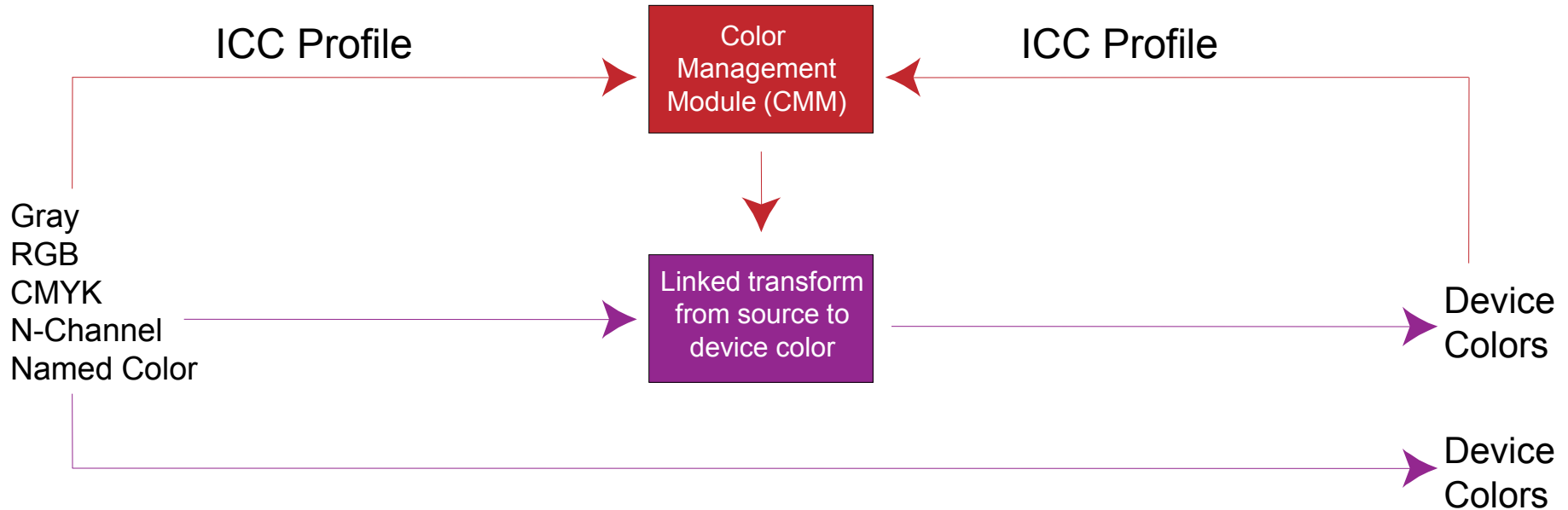
Optimizations that avoid transparency operations in bands that contain no transparency during display list playback.

Performance improvements with softmask handling.

XPSWrite device.

Rework of PDFWrite color to make use of ICC work flow.

Ghostscript Color Flow



Ghostscript Color Architecture

- Easy to interface different CMM with Ghostscript.
- ALL color spaces defined in terms of ICC profiles.
- Linked transformations and internally generated profiles cached.
- Easily accessed manager for ICC profiles.
- Easy to specify default profiles for DeviceGray, DeviceRGB and DeviceCMYK.
- Devices communicate their ICC profiles and have their ICC profile set.
- Operates efficiently in a multithreaded environment.
- Handles named colors with ICC named color profile or proprietary format.
- ICC Color management of Device-N colors or customizable spot handling.
- Includes object type (e.g. image, graphic, text) and rendering intent into the computation of the linked transform. Maintained with transparency.

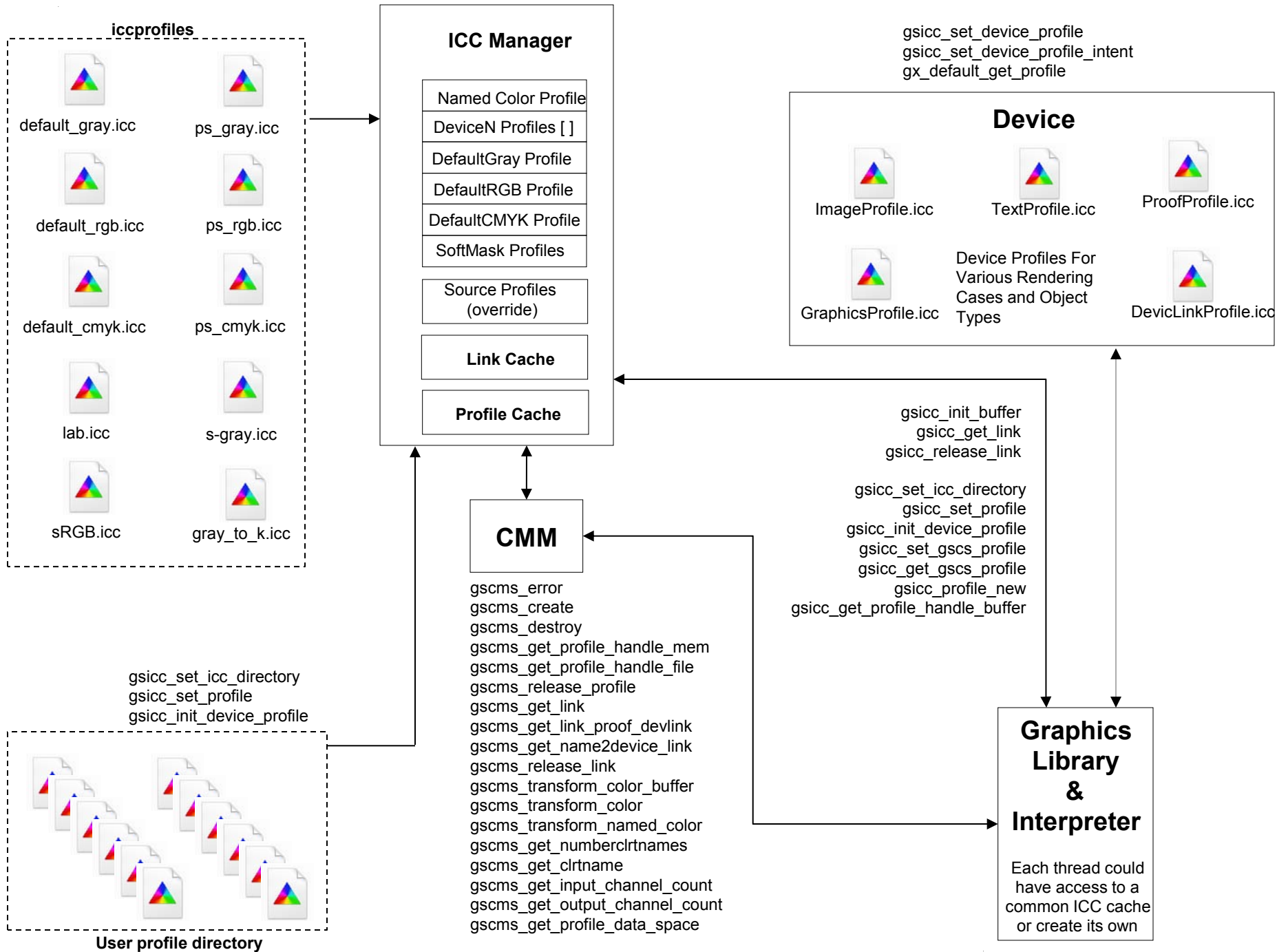
Ghostscript Color Architecture

- Ability to override document embedded ICC profiles with Ghostscript's default ICC profiles.
- Easy to specify unique **source** ICC profiles to use with CMYK and RGB graphic, image and text objects.
- Easy to specify unique **destination** ICC profiles to use with graphic, image and text objects.
- Easy to specify different rendering intents (perceptual, colorimetric, saturation, absolute colorimetric) and black point comp. for graphic, image and text objects.
- Control to force gray source colors to black ink only for devices that support black ink (e.g. CMYK).



Ghostscript Color Architecture

- Make use of PDF output intent ICC profile.
- Use an NCLR ICC output profile when rendering to a separation device.
- Make use of device link ICC profiles for direct mapping of source colors to the device color space.
- Ability to make use of device link ICC profiles for retargeting from SWOP/Fogra standard color space to a specific device color space.



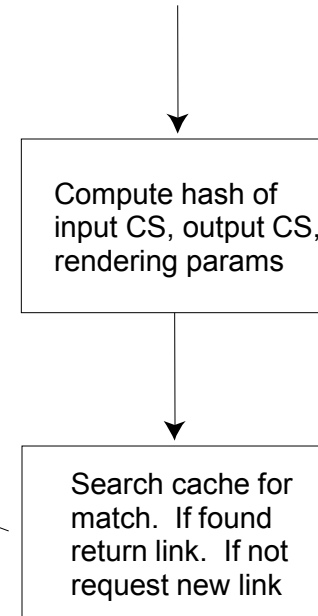
Link Cache

GRAPHICS LIBRARY

```
gsicc_get_link(* pis, *input_colorspace, *output_colorspace, *rendering_params,
memory, include_softproof)
```

Link Cache

Hash Code	Ref Count	Link Structure
Hash Code	Ref Count	Link Structure
Hash Code	Ref Count	Link Structure
Hash Code	Ref Count	Link Structure
.	.	.
.	.	.
.	.	.
.	.	.
Hash Code	Ref Count	Link Structure



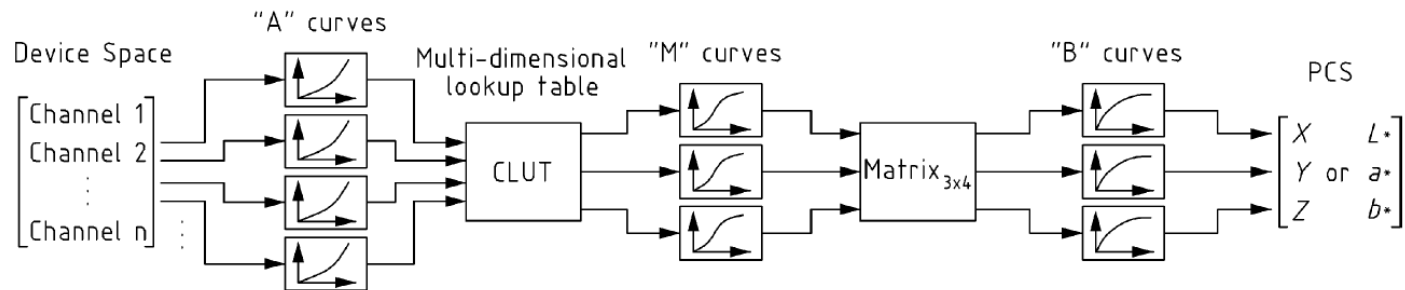
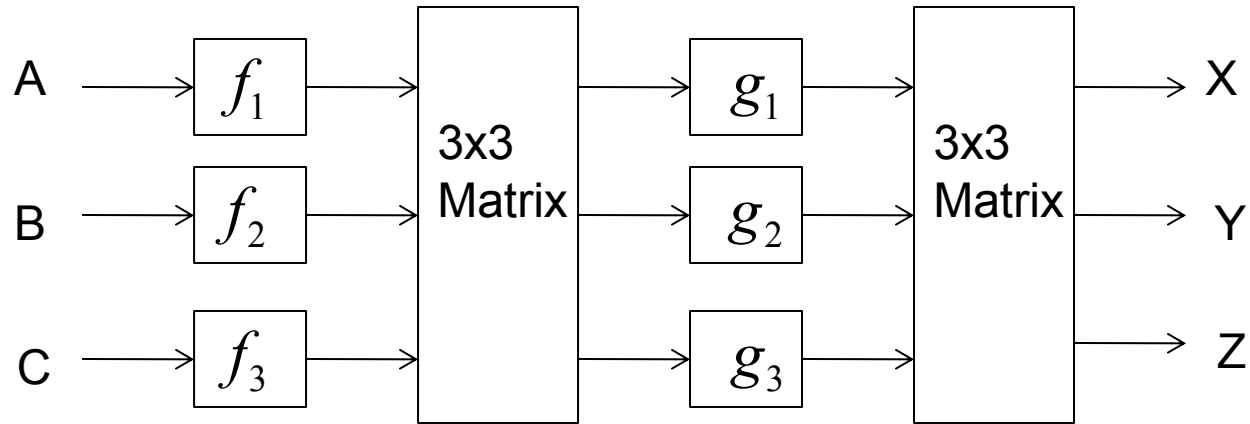
Link entries are reference counted.

Links are only released if we are at maximum number (or memory), new request is made and a Ref Count is one.

Conversion of PS and PDF Color Spaces

- PS and PDF CIE color spaces are converted to ICC forms that the CMM can handle.
- PS mappings are all 1-way. Device to CIEXYZ or CIEXYZ to Device.
- Procedural mappings are sampled.
- Because of the multiple matrix operations and procedural mappings, some PS color spaces that do not include MLUTs will give rise to ICC profiles that do include MLUTs.

Example PS CIEABC



Profile Cache

- Ghostscript creates ICC profiles from PDF and PS CIE colorspace definitions (e.g. CalRGB, CIEABC, CIEDEFG)
- To avoid repeated creations, these profiles are cached based upon a hash code that is related to the resource ID.
- Cache is designed such that MRU item is at the top of the list.
- Profiles are only released if we are at maximum number (or memory), new request is made and a reference count is one.

Device N color spaces (PDF and PS)

- For Device N output, very simple to provide capability for N-color ICC profile.
- Many desire to have CM with CMYK and to pass additional spot colors unmolested.
- For DeviceN input color, XPS requires ICC profile. PDF and PS use an alternate tint transform.
- Architecture provides capability to define N-color ICC profile for DeviceN input colors to replace the alternate tint transform if desired.
- Named color custom support is possible for DeviceN source colors. Example implementation is given in `gs\toolbin\color\named_color` folder

Current Color Command Line Interface

Source Default Profiles

-sDefaultGrayProfile = my_gray_profile.icc
-sDefaultRGBProfile = my_rgb_profile.icc
-sDefaultCMYKProfile = my_cmyk_profile.icc
-sDeviceNProfile = my_devicen.icc
-sNamedProfile = my_namedcolor_profile.icc

Device Profile

-sOutputICCPProfile = my_device_profile.icc

ICC Search Directory

-sICCPProfilesDir = c:/my_iccprofiles/

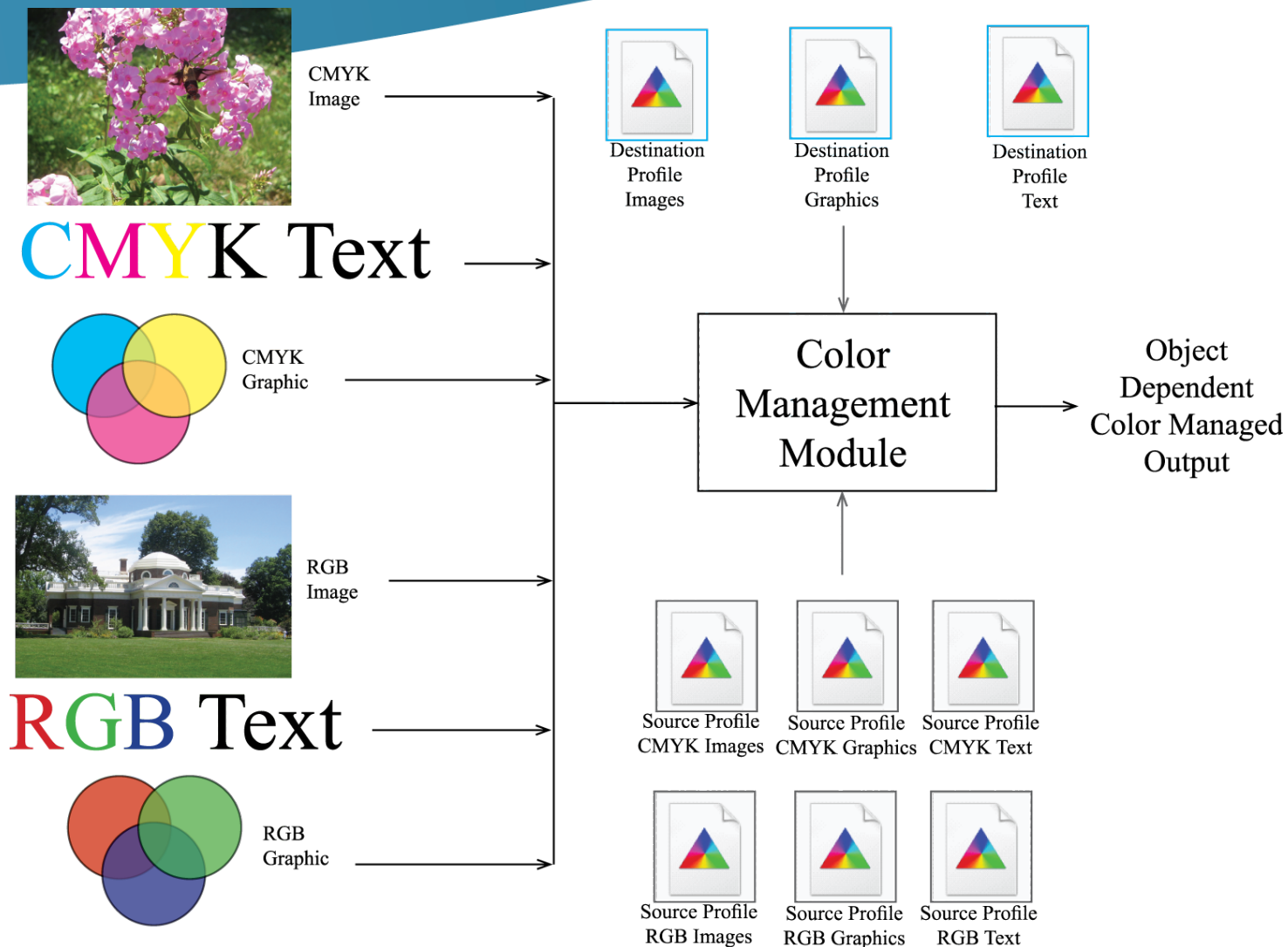
Current Color Command Line Interface

Other Settings

- sProofProfile = my_proof_profile.icc
- sDeviceLinkProfile = my_link_profile.icc
- dRenderIntent = intent (0, 1, 2, 3)
- dOverrideICC = true/false
- dDeviceGrayToK = true/false
- dUseFastColor = true/false
- dBlackPtComp = 0 / 1
- dKPreserve = 0 / 1 / 2
- dSimulateOverprint = true/false
- dUsePDFX3Prole = int

- sICCTransformColors = "Cyan, Magenta, Yellow, Black, Orange, Violet"

Object Dependent Color Management



Object Dependent Color Management Source Profiles

Source object dependent control achieved through the command line Specification:

`-sSourceObjectICC` = filename

Contents of this file define what source profiles and settings should be used with what objects

Key	Profile	Intent	BlackPtComp	Override	BlackPreserve
Graphic CMYK	cmyk_src_graphic.icc	0	1	0	0
Image CMYK	cmyk_src_image.icc	0	1	0	0
Text CMYK	cmyk_src_text.icc	0	1	0	0
Graphic RGB	rgb_source_graphic.icc	0	1	0	
Image RGB	rgb_source_image.icc	0	1	0	
Text RGB	rgb_source_text.icc	0	1	0	

Object Dependent Color Management Destination Profiles

Destination object dependent control achieved through the command line

-sTextICCProfile	=	my_device_text_profile.icc
-sGraphicICCProfile	=	my_device_graphic_profile.icc
-sImageICCProfile	=	my_device_image_profile.icc
-dTextIntent	=	intent (0, 1, 2, 3)
-dGraphicIntent	=	intent (0, 1, 2, 3)
-dImageIntent	=	intent (0, 1, 2, 3)
-sTextBlackPt	=	0/1
-sGraphicBlackPt	=	0/1
-sImageBlackPt	=	0/1
-sTextKPreserve	=	0/1/2
-sGraphicKPreserve	=	0/1/2
-sImageKPreserve	=	0/1/2

Example: Object Dependent CM Default Profiles



RGB Image



RGB Graphic

RGB TEXT

Source file includes RGB and CMYK
Images, graphics and text.



CMYK Image



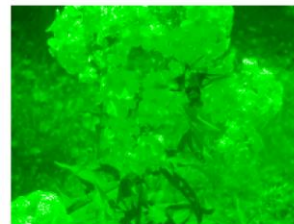
CMYK Graphic

CMYK TEXT

Example: Object Dependent CM Source Profiles Vary

In this case, different ICC profiles were specified to be used with RGB and CMYK graphic, image, and text objects via the Ghostscript command line with `-sSourceObjectICC = filename`.

Graphic_CMYK	cmyk_src_cyan.icc	0 1 0 0
Image_CMYK	cmyk_src_magenta.icc	0 1 0 0
Text_CMYK	cmyk_src_yellow.icc	0 1 0 0
Graphic_RGB	rgb_source_red.icc	0 1 0
Image_RGB	rgb_source_green.icc	0 1 0
Text_RGB	rgb_source_blue.icc	0 1 0

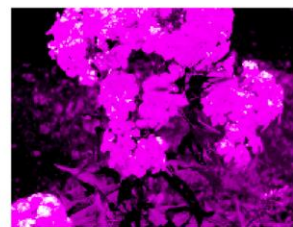


RGB Image



RGB Graphic

RGB TEXT



CMYK Image



CMYK Graphic

CMYK TEXT

Example: Object Dependent CM Source CMYK rendering intent varies

In this case, a special source ICC profile for CMYK objects was specified via the Ghostscript command line. The profile was designed to give radically different results in different rendering intents.

Different rendering intents used for CMYK graphics, images and text

Graphic_CMYK	cmyk_src_renderintent.icc	0	1	0	0
Image_CMYK	cmyk_src_renderintent.icc	1	1	0	0
Text_CMYK	cmyk_src_renderintent.icc	2	1	0	0



RGB Image



RGB Graphic

RGB TEXT



CMYK Image



CMYK Graphic

CMYK TEXT

Example: Object Dependent CM Destination Profile varies

Different destination profiles
specified for different objects

- sGraphicICCProle = yellow_output.icc
- sImageICCProle = magenta_output.icc
- sTextICCProle = cyan_output.icc



RGB Image



RGB Graphic

RGB TEXT



CMYK Image



CMYK Graphic

CMYK TEXT

Example: Object Dependent CM Destination Intent varies

In this case, a special source ICC profile for CMYK objects was specified via the Ghostscript command line.

Different rendering intents used for graphics, images and text

```
-sGraphicICCProle = cmyk_des_renderintent.icc  
-sImageICCProle  = cmyk_des_renderintent.icc  
-sTextICCProle   = cmyk_des_renderintent.icc  
-dImageIntent    = 0  
-dGraphicIntent  = 1  
-dTextIntent     = 2
```



RGB Image



RGB Graphic

RGB TEXT



CMYK Image



CMYK Graphic

CMYK TEXT

Proof and DeviceLink ICC Profile Usage

Two situations:

- 1) Can I print (or display) on device B what my output will look like if I were to print on device A?

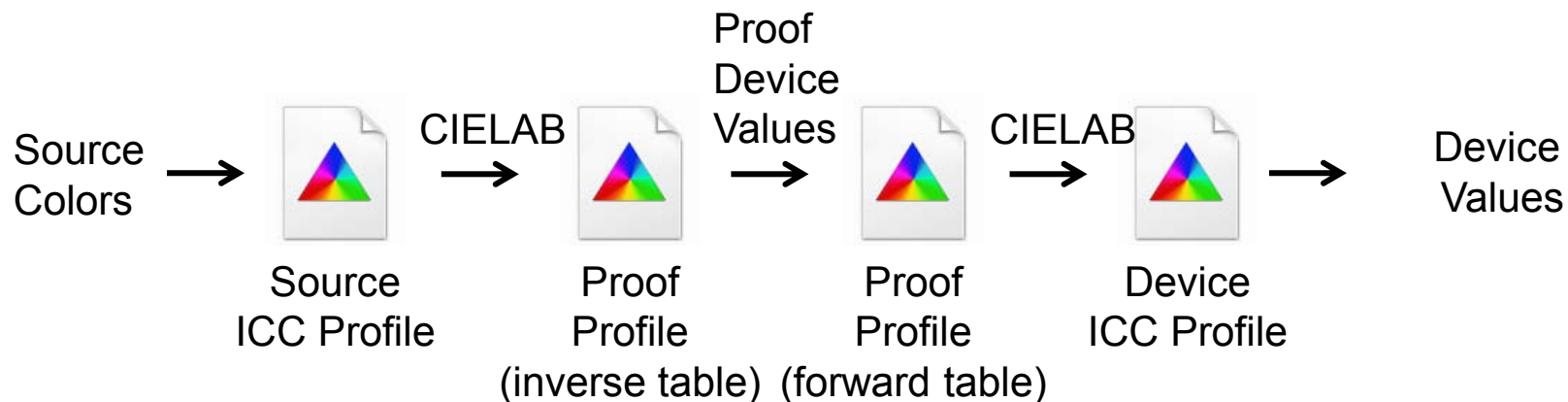
Use a proofing profile.

- 2) Can I map my output to a common standard space (e.g. Forgra39) and then perform a device link transform to my actual device values?

Use a device-link profile.

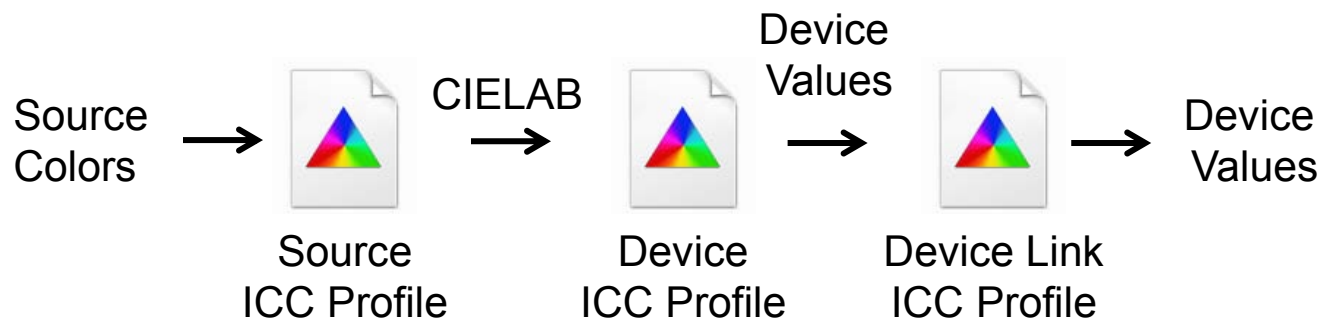
Proof and DeviceLink ICC Profile Usage

Proof Profile Only Case:



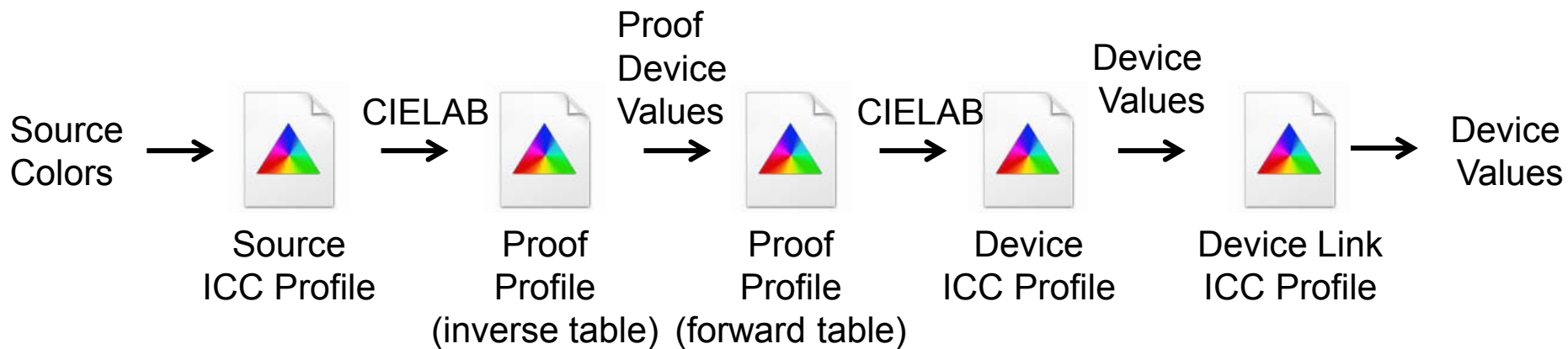
Proof and DeviceLink ICC Profile Usage

Device Link Profile Only Case:



Proof and DeviceLink ICC Profile Usage

Both proofing and device-link profile.



New Case: DeviceLink ICC Profile Usage

Use Device Link Profile Directly to output: Specified with `-sSourceObjectICC`

Graphic_RGB linkRGBtoCMYK.icc 0 1 0

Note output can be DeviceN.



Note that if `-sDeviceLinkProfile` is also used then you have:



Bug Tracking

<http://bugs.ghostscript.com/>

Bugzilla – Main Page version 4.2.5

[Home](#) | [New](#) | [Browse](#) | [Search](#) | [?] | [Reports](#) | [Preferences](#)
| [Administration](#) | [Log out michael.vrhel@artifex.com](#)

Welcome to Bugzilla



[File a Bug](#)



[Search](#)



[User Preferences](#)

[Quick Search help](#)

[Bugzilla User's Guide](#) | [Release Notes](#)

[Home](#) | [New](#) | [Browse](#) | [Search](#) | [?] | [Reports](#) | [Preferences](#)
| [Administration](#) | [Log out michael.vrhel@artifex.com](#)

[My Bugs](#) | [All Mine](#) | [My Closed](#) | [My Customer Bugs](#) | [regresssions_mjv](#)
[customer bugs](#)

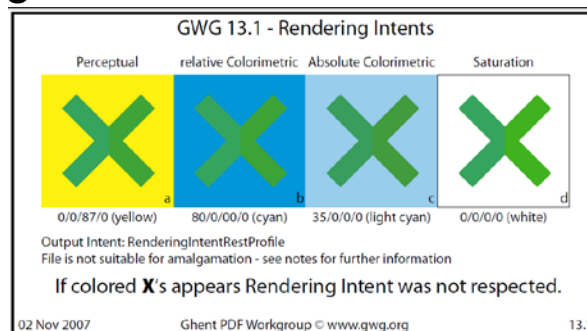
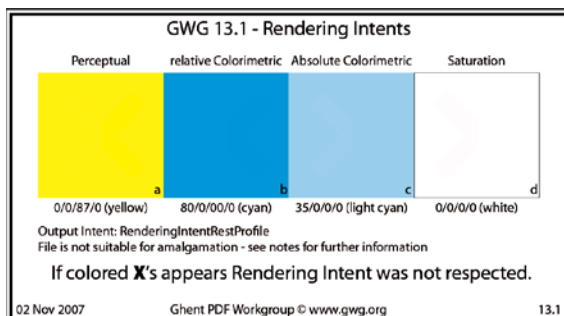
No Significant CUPs device issues

Search in Bugzilla reveals mainly issues from recent new automated tests using valgrind and fuzzing methods

693766	Ghostscr	Valgrind	h@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in do_validate_chunk	2013-03-31
693777	Ghostscr	Valgrind	@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in gs_image_class_1_simple	2013-03-31
693781	Ghostscr	Valgrind	el@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in gx_color_frac_map	2013-03-31
693782	Ghostscr	Valgrind	@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in gx_dc_default_fill_masked	2013-03-31
693784	Ghostscr	Valgrind	@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in gx_default_copy_color	2013-03-31
693787	Ghostscr	Valgrind	el@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in gx_restrict_ICC	2013-03-31
693791	Ghostscr	Valgrind	h@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in ialloc_validate_ref_packed	2013-03-31
693794	Ghostscr	Valgrind	el@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in image_render_interpolate_icc	2013-03-31
693795	Ghostscr	Valgrind	@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in image_render_mono	2013-03-31
693809	Ghostscr	Valgrind	@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in orig_sqrt	2013-03-31
693812	Ghostscr	Valgrind	el@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in pdf14_encode_color	2013-03-31
693819	Ghostscr	Valgrind	el@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in threshold_16_SSE	2013-03-31
693820	Ghostscr	Valgrind	el@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in threshold_16_SSE_unaligned	2013-03-31
693822	Ghostscr	Valgrind	@artifex.com	CONF	---	Valgrind reports uninitialised value(s) in width_is_thin	2013-03-31
693968	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in cio_bytein	Thu 01:26
693971	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in cmd_compress_bitmap	Thu 01:32
693972	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in cmd_drawing_color_usage	Thu 01:34
693973	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in cmd_write_band	Thu 01:36
693974	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in cmd_write_pseudo_band	Thu 01:38
693983	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in do_validate_chunk	Thu 01:54
693984	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in Draw_Sweep	Thu 01:56
693985	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in dwt_decode	Thu 01:58
693986	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in Finalize_Profile_Table	Thu 01:59
693988	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in FT_Done_Size	Thu 02:03
693994	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in gs_scan_token	Thu 02:14
693998	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in gsicc_get_device_profile_comps	Thu 02:21
694006	Ghostscr	Fuzzing	tifex.com	CONF	---	Seg faults found by fuzzing in gx_forward_dev_spec_op	Thu 02:35

PDF Output Rendering Intent

GS now supports PDF Output Rendering Intent usage Passes all Ghent tests and these are included in the testing suite.



OutputIntents array (Optional; PDF 1.4) An array of output intent dictionaries describing the color characteristics of output devices on which the document might be rendered (see “Output Intents” on page 970).

PDF Output Rendering Intent

`-dUsePDFX3Profile = #`

Where # defines which output intent to use in the order that they occur in the document. If no number specified, first one encountered is used.

If no profile is present in the intent dictionary, a warning is displayed and the rendering intent is ignored.

If the output intent ICC profile does not match the process color model of the output device, then the output intent ICC profile is used as a proofing profile.

Planar Separation Devices

Most devices make use of a memory device to buffer the rendered page.

Until recently, Ghostscript was primarily set up for use with chunky memory with the largest chunky pixel being 64 bits.

This presented a limitation for Separation devices with a large number of spot colors.



64 bit word with CMYK + 4 spots

Planar Separation Devices

One approach to solve this was to use a compressed color encoding scheme.

Since certain combinations are more likely to occur.

For example a pure 100% spot with no other colorants is going to be common in label printing.

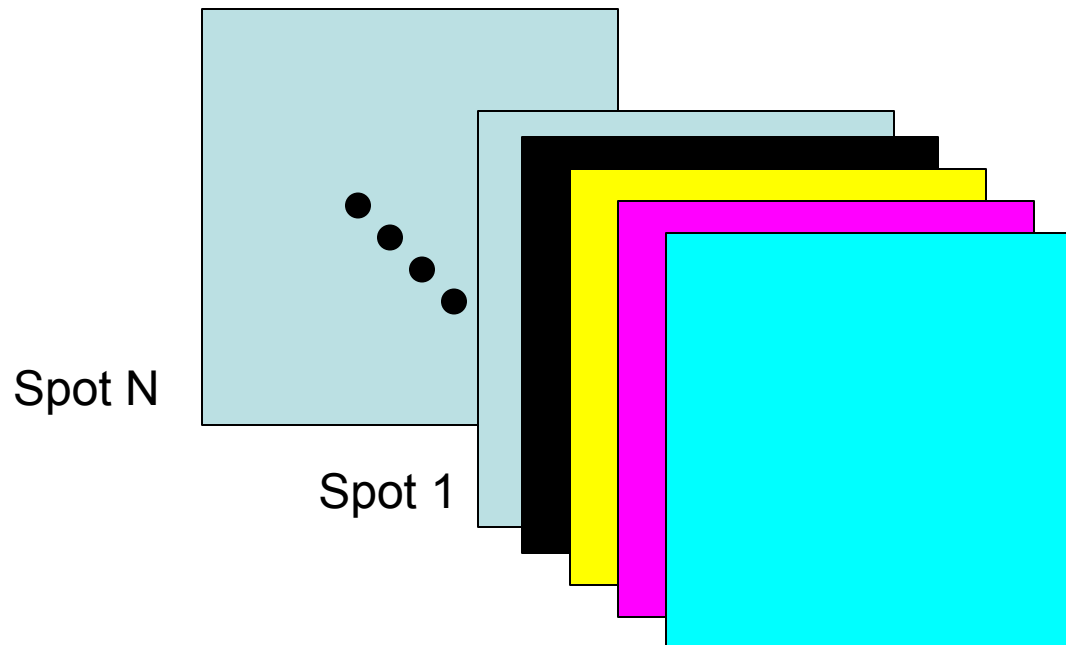
In the presence of transparency and shadings **with** multiple colorants this approach begins to break down.

Planar Separation Devices

Solution is to go to a planar memory memory model.

Advantage for laser print applications.

Now fully implemented and regression tested with every commit.





MuPDF



What is MuPDF?



- **A core set of libraries**
 - + Entirely written in C
 - + Very portable
 - We've done: Windows/Linux/MacOS/iOS/Android
 - Third parties: BB10/QNX/others
- Various example tools that use these libraries:
 - + Simple viewers for Linux/Android/MacOS/iOS/Windows.
 - + Command line tools for rendering PDF pages.
 - + Command line tools for manipulating PDF files
 - page extraction
 - decompression
 - repair
 - resource extraction
- Licensing the same as Ghostscript. Dual AGPL/Proprietary licensed. Artifex owns the copyright.

Why MuPDF when we have Ghostscript?



For printing on large machines - use Ghostscript.

- * Postscript
- * PCL
- * Spot colors
- * Extreme level of color management control
- * Massive range of output devices

Why MuPDF when we have Ghostscript?



For screen use or embedded devices - use MuPDF.

*** Fast**

- + PDF Parser in C, not Postscript.
- + AA Rendering designed in from the ground up.

*** Small**

- + Much smaller ROM footprint.

*** Simple**

- + No complex garbage collector to maintain
- + Small set of dependent libraries
- + Simpler to port

*** Interactive features**

- + More suitable for building viewers
- + Searching
- + Zooming
- + Form filling
- + Transitions

Features of MuPDF



"Complete" PDF support

- + Transparency
- + Patterns/Shadings
- + Fonts (all kinds)
- + Image formats
- + Decryption
- + Interaction (more later)

Features of MuPDF



Not just "PDF" - Other formats too:

- + XPS
- + CBZ/JPEG/PNG
- + Extensible system

"Device" interface

- + Separates interpretation from rendering
- + Allows display lists
 - interpret page once, render many times at different zooms
- + Allows format conversions (more later)

Clever memory management

- + Caching of objects (both raw and decoded)
- + Memory scavenging (throw objects away just in time)

Optional Multi-core/threading support

- + Not tied to any one threading implementation
 - All we need is locks
- + Interpretation happens on 1 thread
- + Rendering can happen on many.
 - Thumbnails rendered in the background.
 - Banded rendering of pages.
 - Resources decoded on one thread can be used by others.

Recent changes: Interactivity



Form filling

- + Javascript for validation
 - Not tied to any one javascript implementation
 - Thin veneer to Googles 'v8' engine supplied
- + Ability to save files back with data in them.

Google Cloud Print support added to Android app

Reflow View

- + Pages extracted to HTML (text and images).
- + Rudimentary layout detection (tables, indents, etc.)
- + Still a work in progress.

Digital Signatures

- + Verify signatures
- + Sign documents
- + Re-sign documents after form filling
- + Still a work in progress.

Submission of filled in forms.

- + Several different ways of doing this.
 - Send the whole filled in file.
 - Extract data as XML and send that.

Improve page extraction for reflow.

- + Improve column detection
- + Reorder lines
- + Spot captions on images
- + Preserve "line art" areas of the page as images

Output devices for format conversion.

- + PDF output (prototype code exists)
- + SVG output (prototype code exists)
- + Other possibilities include XPS or PCL.

More input devices

- + SVG seems most likely.

Color Management

- + Use LCMS to enable color management throughout.



Get Involved



If any of this sounds interesting, we'd love to hear from you.

Our development direction is driven by customers and users.

Contributions are always welcome. We have a bug bounty scheme.

Repository located at

`git://git.ghostscript.com/mupdf.git`

MuPDF discussions on IRC freenode #ghostscript channel

Additional information at www.mupdf.com



Thank you for your attention!



michael.vrhel (at) artifex.com