

Ghostscript and MuPDF Status OpenPrinting Summit May 2020

Michael Vrhel, Ph.D.
Artifex Software Inc.
Novato CA



Outline



Ghostscript

MuPDF

Ghostscript – Changes since last meeting

MuPDF – Changes since last meeting

Little CMS2MT

MuJS

MuPDF WebAssembly

Current and Future Work

CUPS Device

Ghostscript



Ghostscript is a document conversion and rendering engine.

Converts between and renders PDF, PS, PCL, PCL-XL, XPS

Dual license GNU AGPLv3 / Commercial

Source and documentation available at www.ghostscript.com

MuPDF



Open source software framework for viewing and converting PDF, XPS, and e-book documents.

Designed toward mobile environment use.

Dual license GNU AGPLv3 / Commercial

Source and documentation available at www.mupdf.com

MuPDF



Command Line Tools

mutool draw

Primarily used for rendering a document to image files.

mutool convert

For converting documents into other formats.

mutool trace

Debugging tool used for printing a trace of the graphics device calls on a page.

Command Line Tools

mutool show

A tool for displaying the internal objects in a PDF file.

mutool extract

Extract images and embedded font resources.

mutool clean

Rewrite PDF file. Used to fix broken files, or to make a PDF file human editable.

MuPDF



Command Line Tools

mutool merge

Merge pages from multiple input files into a new PDF.

mutool create

Create a new PDF file from a text file with graphics commands.

mutool run

A tool for running Javascript programs with access to the MuPDF library functions.

MuPDF



JavaScript

Examples in docs/examples:

pdf-merge.js

pdf-portfolio.js

pdf-create.js

and more...

Example:

```
mutool run pdf-merge.js output.pdf input1.pdf input2.pdf ...
```


Changes to GS since last meeting



Release 9.50 October 2019

- Redesign of file access for security
- File access controls enabled by default. In order to run Ghostscript without these controls, use `-dNOSAFER`
- Postscript non-standard and undocumented operators no longer accessible
- 16 bit per component capability in transparency device
- Implementation of compatible overprint mode

Changes to GS since last meeting



Release 9.52 March 2020

- Implementation of fillstroke device method (fixes issues with transparency in those operations)
- Zero Coverity issues achieved (will be maintained moving forward)
- Introduction of SIMD acceleration methods for commercial licenses (Transparency blending, interpolation, halftoning, color conversion)
- Moved to Visual Studio 2019 solution

Changes to MuPDF since last meeting



Release 1.15: (May 2019)

- WebAssembly build target and examples
- Improved forms API in both C and Java bindings.
- Improved forms JavaScript support.
- Create appearance streams for more form field types.
- Improved text extraction from LaTeX documents with math symbols.
- Support CBZ and CBT files larger than 2Gb.
- Show table of contents for FB2 and XHTML documents.
- Show embedded raster images in SVG and XHTML documents.
- Show FB2 cover page.
- Add inhibit-spaces option to stext device to turn off missing space detection.
- Simplified fz_try/fz_always/fz_catch macros.

Changes to MuPDF since last meeting



Release 1.15: (May 2019)

mupdf-gl improvements:

- Automatically open annotation editor when selecting an annotation
- Full page color tinting option in mupdf-gl
- Show/hide table of contents sections
- Toggle spot color mode with 'e'

Changes to MuPDF since last meeting

Release 1.15: (May 2019)

mutool improvements:

mutool show \$PDF js

show document level javascript.

mutool clean -A

create appearance streams when missing.

mutool clean -AA

recreate all appearance streams.

mutool run docs/examples/portfolio.js

extract embedded files from PDF document.

mutool run docs/examples/pdf-dejpx.js

decompress JPEG2000 images in PDF document.

mutool run docs/examples/fix-s22pdf.js

fix fonts in documents generated by S22PDF.



Changes to MuPDF since last meeting

Release 1.15: (May 2019)

Various cleanups:

- Removed TGA output support. Use one of the more common and useful PAM, PNM, PNG, or TIFF formats instead.
- Removed support for PDF portfolios. This work-in-progress feature was never completed, and only worked for a small subset of files.
- Removed support for progressive loading. This rarely used feature added a lot of complexity and was an unending source of bugs and strange behavior. Removing this feature has allowed us to clean up and simplify large amounts of code, and fix dozens of bugs in one fell swoop.



Changes to MuPDF since last meeting

Release 1.16: (July 2019)

API:

- Major overhaul of color management architecture
- Improved functions to verify/sign PDF documents
- Number tree accessor function `pdf_lookup_number`
- Parse and handle more options for PWG output
- Removed obsolete gproof document type
- User callbacks for warning and error messages



Changes to MuPDF since last meeting

Release 1.16: (July 2019)

epub:

- Changed default page size to A5.
- Draw embedded SVG documents in EPUB/XHTML.
- New Noto fonts covering more unicode scripts.
- Support small-caps font-variant.

pdf:

- Add Redact annotation type and function to apply redactions.
- Add/remove/change encryption and password when saving PDF files.
- Improvements to text handling in PDF filter processor.
- New and improved progressive loading.

svg:

- Draw external images in SVG documents.



Changes to MuPDF since last meeting

Release 1.16: (July 2019)



mutool show:

Add 'form' selector to list PDF form fields.

mutool sign:

Sign PDF documents from the command line.

Verify signatures in PDF documents from the command line.

viewer:

Option to save a script of user actions for replay with mutool run.

Runtime toggle for ICC and spot rendering.

Tooltip when hovering over a link in mupdf-gl.

Changes to MuPDF since last meeting

Release 1.17: (Any day now)

api:

- Improved accessors for markup/ink/polygon annotation data.
- Chapter based API for faster EPUB loading.
- Add more documentation to header files.
- Improved digital signature signing and verification.
- Validate changes in a signed PDF file.

build:

- Moved windows build to VS2019 solution.

pdf:

- Redaction now works on images and links as well as text.
- Greek, Cyrillic, Chinese, Japanese, and Korean scripts in forms and annotations.
- File attachment annotations.
- Use CCITT Fax compression for 1-bit images when creating PDF files.



Changes to MuPDF since last meeting



Release 1.17: (Any day now)

epub:

- More forgiving XHTML parsing.

- Accelerator files to cache chapter data for faster EPUB loading.

- Optimized memory use.

mutool run:

- Edit Markup, Ink, and Polygon annotation data.

- Fill out form fields.

viewer:

- Ask for confirmation before closing a PDF with unsaved changes.

- Embed and extract file attachment annotations.

Little CMS2MT

We continue to use a fork of Little CMS2 that is thread safe.

Fork is currently available with git checkout of Ghostscript.

We bring in any bug fixes applied to Little CMS2 .

Possible to make use of 3rd party plug-ins.

Developed SSE4.2, AVX2, NEON plug-in for tetrahedral interpolation with cmyk, rgb, or gray output.
(Commercial license only)

MuJS



MuJS is a library, written in C.
MuJS has no notion of a main program:
it only works embedded in a host client program.
The host program can invoke functions to execute
Javascript code, read and write Javascript variables,
and register C functions to be called by Javascript.

Implements ECMAScript ECMA-262

Open source under ISC license

<https://opensource.org/licenses/ISC>

<http://git.ghostscript.com/?p=mujs.git;a=summary>

git.ghostscript.com Git · × +

git.ghostscript.com/?p=mujs.git;a=summary

[projects](#) / [mujs.git](#) / [summary](#)

[summary](#) | [shortlog](#) | [log](#) | [commit](#) | [commitdiff](#) | [tree](#)

description MuJS: a Javascript interpreter
owner Artifex Software, Inc.
last change Tue, 17 Mar 2020 06:10:18 -0700 (14:10 +0100)
URL <git://git.ghostscript.com/mujs.git>
<http://git.ghostscript.com/mujs.git>

shortlog

2020-03-17	Tor Andersson	Fix potential memory corruption when jsV_newmemstring... master 1.0.7	commit commitdiff tree
2020-03-17	Tor Andersson	Return "[object Iterator]" when running toString on...	commit commitdiff tree
2020-03-17	Tor Andersson	Document the expected behavior of js_Put correctly.	commit commitdiff tree
2020-03-17	Tor Andersson	Note that js_Report callback must not throw an exception.	commit commitdiff tree
2020-02-19	Tor Andersson	Check for empty string in js_isarrayindex.	commit commitdiff tree
2020-01-23	Tor Andersson	Tweak default recursion limit.	commit commitdiff tree
2020-01-23	Tor Andersson	Fix enumerability of Error and Function properties.	commit commitdiff tree
2020-01-20	Tor Andersson	Check for leading zero in js_isarrayindex that caused...	commit commitdiff tree
2020-01-10	Tor Andersson	Bug 698496: Handle leap years correctly in the Date...	commit commitdiff tree
2020-01-10	Tor Andersson	Bug 701355: Fix separate compilation issues that have...	commit commitdiff tree
2020-01-02	Tor Andersson	Bug 701886: Always create new scope for eval().	commit commitdiff tree
2020-01-02	Tor Andersson	Update COPYING copyright year.	commit commitdiff tree
2020-01-02	Tor Andersson	Issue 117: Skip first line if it starts with a shebang...	commit commitdiff tree
2020-01-02	Tor Andersson	Issue 128: Support property list as replacer argument...	commit commitdiff tree
2020-01-02	isRyven	Fix date setMonth and setUTCMonth processed wrong optio...	commit commitdiff tree
2020-01-02	isRyven	Issue 126: accept String and Number objects as space...	commit commitdiff tree

MuPDF WebAssembly

Binary instruction format for a stack-based virtual machine.



Wasm is designed as a portable target for compilation enabling deployment on the web for client and server applications.

<code>mupdf.openDocument(filename)</code>	Open a document and return a handle.
<code>mupdf.freeDocument(doc)</code>	Free a document and its associated resources.
<code>mupdf.documentTitle(doc)</code>	Return the document title as a string.
<code>mupdf.documentOutline(doc)</code>	Return element containing the table of contents formatted as an unordered HTML list with links to pages
<code>mupdf.countPages(doc)</code>	Return the number of pages in the document.
<code>mupdf.pageWidth(doc, page, dpi)</code>	Return the width of a page.
<code>mupdf.pageHeight(doc, page, dpi)</code>	Return the height of a page.
<code>mupdf.drawPageAsPNG(doc, page, dpi)</code>	Render the page and return a PNG image formatted as a data URI.
<code>mupdf.pageLinks(doc, page, dpi)</code>	Retrieve an HTML string describing the links on a page.
<code>mupdf.drawPageAsSVG(doc, page)</code>	Return a string with the contents of the page in SVG format.
<code>mupdf.drawPageAsHTML(doc, page)</code>	Return a string with the contents of the page in HTML format, using absolute positioned elements.

<https://ghostscript.com/~tor/wasm/view.html>

Current/Future Work



New PDF interpreter in Ghostscript. Significant progress. See pdfi branch in repository

Python and other language support for MuPDF

Overprint emulation for RGB devices in Ghostscript

Support for full black only text output in Ghostscript

Device API cleanup in Ghostscript color methods

Image interpreters and language switching in Ghostscript

CUPS Device



702132 - Ghostscript dr x + v

https://bugs.ghostscript.com/show_bug.cgi?id=702132

Bugzilla - Bug 702132 Ghostscript does not correctly create CMY CUPS Raster and often crashes

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

Bug List: (2 of 8) [First](#) [Last](#) [Prev](#) [Next](#) [Show last search results](#)

Bug 702132 - Ghostscript does not correctly create CMY CUPS Raster and often crashes ([edit](#))

Status: UNCONFIRMED ([edit](#))

Alias: None ([edit](#))

Product: Ghostscript

Component: CUPS driver ([show other bugs](#))

Version: 9.50

Hardware: PC Linux

Importance: P4 normal

Assignee: [Till Kamppeter](#) ([edit](#)) ([take](#))

QA Contact: [Bug traffic](#) ([edit](#)) ([take](#))

URL:

Keywords:

Personal Tags:

Depends on:

Blocks:

Reported: 2020-02-17 17:39 UTC by [Till Kamppeter](#)

Modified: 2020-02-17 17:39 UTC ([History](#))

CC List: Add me to CC list
2 users ([edit](#))

Ignore Bug Mail: (never email me about this bug)

See Also: ([add](#))

Customer:

Word Size: ---

Orig. Est.: **Current Est.:** **Hours Worked:** **Hours Left:** **%Complete:** **Gain:** **Deadline:** **Only users in all of the selected groups can view this bug:**

Code Security/Analysis Methods



Fuzzing of test files used to detect simple faults

Coverity : <https://scan.coverity.com/projects/ghostpd> (dereferences of NULL pointers, use of uninitialized data, memory corruptions, buffer overruns, control flow issues, incorrect expressions, unsafe signed values)

Coverage tests run periodically: <https://ghostscript.com/coverage/>

Various compilers used and warning report provided with every commit (gcc, clang)

Address Sanitizer: Testing for buffer overflows, dangling pointer overflows

Valgrind: Testing for buffer overflows, use of uninitialized memory/variables

Memento: Memory leak/corruption analyzer and “Memory Squeezing”. Part of Ghostscript build memento.h/c

More Information



Repositories located at
[git://git.ghostscript.com](https://git.ghostscript.com)

Ghostscript discussions on IRC freenode #ghostscript channel
MuPDF discussions on IRC freenode #mupdf channel

Bug reports
bugs.ghostscript.com

Additional information at www.mupdf.com www.ghostscript.com