

Report PDF Emitter

Version: Draft 2

Abstract

This document describes the features supported by the new PDF emitter in BIRT 2.0.

Document Revisions

Version	Date	Description of Changes
Draft 2	8/12/2005	Updated by Stanley Wang
Draft 1	8/11//2005	Initial draft.

Contents

1. Introduction.....	1
2. Supported Features in PDF Library	2
3. Supported Features in PDF Emitter	3
3.1 ROM Support	3
3.2 Pagination	3
3.3 Hyperlinks	3
3.4 Multiple languages	3
3.5 Performance	4
3.6 Image Formats.....	4
3.7 Document Map.....	4
3.8 Compression.....	4
3.9 Reportlet PDF	4
3.10 Customization	4
4. Related Bugzilla Entries.....	5

1. Introduction

In BIRT 1.0, the PDF generation support is limited by the capabilities of FOP. Such limitations include inadequate performance, large memory footprint, text overlap, and limited support for line, border, font and image formats.

BIRT 2.0 intends to address these issues by building a PDF emitter independent of FOP. Such an independent solution includes two components: a PDF library and an emitter.

2. Supported Features in PDF Library

The PDF library supports a number of basic and advanced PDF features. The criteria to determine whether a feature is included in the library are based on the sole need of PDF emitter, thus certain features deemed important by Adobe may not be included.

The supported PDF features include:

1. Text and drawing objects. They are fundamental building blocks in a PDF document.
2. Image object. Support multiple encodings, such as raw bitmap, Flate encoding, run length encoding and zip compression.
3. Document map (for supporting TOC).
4. Supports RGB and CMYK color spaces.
5. Hyperlink.
6. Flexible page size. A PDF file can have first page in US Letter and second page in A4.
7. True-type font embedding and subsetting.
8. Content compression.
9. Customization, such as changing retribution of the file, and insertion of a block of independent content (for example EPS).
10. Fillable forms. (Not in release 2.0)
11. Linearization. (Not in release 2.0)
12. Security features, i.e., password protection, digital signature, and encryption.

Linearization is not in release 2.0 because it is less useful in reporting environment because it requires knowing the file size beforehand.

The PDF library must support creation of large PDF file efficiently. If the host application sequentially writes out PDF objects, the memory should be bound by the size of the largest object, and should not be dependent on the actual file size.¹

BIRT 2.0 may use other open-source PDF library if such a library supports (most of the) features listed here, and the license terms are compatible to Eclipse license terms. If this is the case, some of the features listed as beyond 2.0 may be supported in BIRT 2.0.

If a PDF library is developed independently, the library would expose a set of APIs for PDF emitter to use. Such APIs are considered as internal in BIRT 2.0 and are not guaranteed to work seamlessly in environment other than BIRT.

The PDF library should generate PDF files that are compatible with Acrobat reader 5.0.

¹ This does not mean that creating a large table would use the same amount of memory as creating a small table. However, the need for more memory to for large table comes from table rendering, not PDF itself.

3. Supported Features in PDF Emitter

3.1 ROM Support

BIRT PDF emitter supports:

1. All report elements supported in BIRT 2.0.
2. All style properties supported in BIRT 2.0. The behavior for each style property is defined by ROM spec. CSS spec would be another source of reference if a property also appears in CSS style, and ROM does not explicitly specifies the behavior.
3. Page number, page sequence and master page, including multiple-column reports.²

3.2 Pagination

PDF emitter supports hard page breaks that are specified by the page-break-after, page-break-before or page-break-inside properties. In addition, it supports automatic page break in both the vertical and horizontal dimensions. See Report Paging specification (BPS19) for more detail.

The actual pagination and layout of a report depends on the fontmetrics that are used to calculate the size of report content. PDF emitter provides a programming interface for other developers to supply character/text size information. It also supports a configuration file for specifying a fontmetrics file to be used by the emitter. Fontmetrics info may come directly from a true-type font file, or a more compact file (format to be determined).

3.3 Hyperlinks

Hyperlink, whether it is a URL, search URL, or a drill through link, is manifested as an external URL in the PDF file, and is supported. Linking to another report document always links to the HTML format. Linking to another report design always links to generation page of a report. If the target report has no unfilled parameter, the report is run and HTML output is displayed. Otherwise, the report parameter page may be shown first. Links always open in a separate browser window.

Bookmarks in the same file are supported in BIRT 2.0 and links to the same PDF file. Linking to a bookmark in a different file is supported and links to the HTML format of the report.

3.4 Multiple languages

PDF emitter supports characters in multiple code pages appeared in the same string.

PDF emitter supports font embedding and subsetting. The exact behavior for font subsetting and embedding can be controlled by a configuration file.

² Multiple-column report is supported only if the designer and engine supports such reports in BIRT 2.0.

3.5 Performance

Performance for PDF generation should roughly be linear to the content that is to be generated. Perfect linearity is not achievable due to certain level of buffering for layout and pagination. However, such buffering should not be excessive to significantly affect performance. For example, pagination should not require buffering a whole table. Instead, if a table column width is not specified, up to 100 rows of data will be scanned and the column width is determined based on the scanned rows; as another example, if a drop column appears with more than 100 rows, and it needs to be aligned middle vertically, it will be placed on 50th row. Other limitations in the same line may be identified during implementation, and they should not be viewed as defects.

3.6 Image Formats

PDF emitter supports all image formats that are supported in report designer. Some of the formats may not be supported natively by PDF format. In such cases, a conversion to other file formats may happen behind the scene.

Images should in general be stored in a compressed format in the PDF files. It is considered a defect if the PDF file is too large because certain image files are not compressed.

BIRT will bundle an open-source image library for image manipulations and processing, if there is no licensing issues. If this is not the case, the user is required to download such a library before using some image formats.

SVG image is supported, but by first converting to other image format in BIRT 2.0. Interactivity is therefore lost. Future release may support image/chart interactivity in PDF.

3.7 Document Map

BIRT TOC drives the generation of document map, which guides the viewing experience once the PDF file is opened.

A TOC at the beginning of a report (like the Table of Content for a book) is not supported in release 2.0.

3.8 Compression

BIRT 2.0 supports compressing objects appeared in a PDF file. It does not support compression at the file level.

3.9 Reportlet PDF

BIRT 2.0 HTML emitter supports generating a report item independently from the page it appears in. Similar feature is not supported in BIRT 2.0 PDF emitter.

3.10 Customization

Previous sub-sections have covered several customization features, such as customizing fontmetrics source, compression, etc. Other customizations include retribution, embedding black-box object, selection of image format in the PDF file, etc. All the customizations are controlled by a single XML configuration file (format to be defined later).

4. Related Bugzilla Entries

- [91079](#) PDF file can't display Multilanguage correctly except Eng
- [91080](#) The contents overlapped in PDF
- [89160](#) PDF report creation unsuccessful
- [101216](#) setting of property "width" on cell gives no effect for PDF.
- [101217](#) setting of property "marginTop" on grid gives no effect f...
- [101369](#) Preview as PDF gets Adobe Reader error 109
- [101375](#) View in PDF chops off Columns
- [101450](#) Invalid PDF Generated
- [101971](#) Insert image into master page header, can not preview it
- [102241](#) PDF generated using Korean report content displays Korean.
- [103133](#) Apply pre-defined style to a table, "underline" of font doesn't work.
- [103316](#) '#' instead of cyrillic letters in pdf-viewing
- [103665](#) PDF Export Losing Formating Multi Colum Repor
- [104319](#) Image size doesn't work properly when previewed in pdf.
- [104758](#) all the pdf page should be gray if preview in PDF after s...