IEEE-ISTO
Printer Working Group
Portable Document Format: Image-Streamable (PDF/is)

Working Draft
Maturity: Prototype

5 August 2003

Deleted: 1 July 2003

Deleted: 1 July 2003

IEEE-ISTO
Printer Working Group
Portable Document Format: Image-Streamable
(PDF/is)

Working Draft
Maturity Level: Prototype

5 August 2003

Deleted: 1 July 2003

Abstract: This document specifies an application of PDF (Portable Document Format) that has two important properties: First, it is an "image"-based format, and proper rendering of the document is represented by (binary or color) images. Second, the format is suitable for incremental generation and thus it is a "streaming" format. The subset is called "PDF/is", for "PDF Image-Streamable".

PDF/is is formally a subset of PDF 1.4, and is intended to be fully compatible with software that reads PDF 1.4. There are "profiles" of PDF/is, which are distinguished primarily by the methods if image compression and/or techniques employed. The representations of image data employed are specified in the PDF 1.4 language reference [pdf], which in turn describes the PDF representation of image data specified by ITU-T recommendations for black-and-white facsimile ([t.4], [t.6]), ISO/IEG specifications for digital compression and coding of continuous-tone still images [jpeg], and lossy/lossless coding of bi-level images [jbig2].

PDF/is is intended to be useful within the IPPFAX protocol [reference], which is used to provide a synchronous, reliable exchange of image documents between senders and receivers. For this reason, PDF/is also includes an optional security features for digital signaturing.

Deleted: 1 July 2003

This document is available electronically at:
ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030805.pdf,
ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030805.doc

Deleted: 30
Field Code Changed
Deleted: 6
Field Code Changed

A version showing the changes from the previous version is available at:
ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030805-rev.pdf

Deleted: 630
Deleted: 0630
Field Code Changed

The latest version of this specification is available at:
ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-latest.pdf,
ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-latest.doc

For a definition of "Maturity Level" used on the title page, along with any other questions about the Printer Working Group's processes, please see the PWG process document [process].

About the IEEE-ISTO

The IEEE-ISTO is a not-for-profit corporation offering industry groups an innovative and flexible operational forum and support services. The IEEE-ISTO provides a forum not only to develop standards, but also to facilitate activities that support the implementation and acceptance of standards in the marketplace. The organization is affiliated with the IEEE (http://www.ieee.org/) and the IEEE Standards Association (http://standards.ieee.org/).

For additional information regarding the IEEE-ISTO and its industry programs visit http://www.ieee-isto.org.

About the IEEE-ISTO PWG

The Printer Working Group (or PWG) is a Program of the IEEE Industry Standards and Technology Organization (ISTO) with member organizations including printer manufacturers, print server developers, operating system providers, network operating systems providers, network connectivity vendors, and print management application developers. The group is chartered to make printers and the applications and operating systems supporting them work together better. All references to the PWG in this document implicitly mean "The Printer Working Group, a Program of the IEEE ISTO." In order to meet this objective, the PWG will document the results of their work as open standards that define print related protocols, interfaces, procedures and conventions. Printer manufacturers and vendors of printer related software will benefit from the interoperability provided by voluntary conformance to these standards.

In general, a PWG standard is a specification that is stable, well understood, and is technically competent, has multiple, independent and interoperable implementations with substantial operational experience, and enjoys significant public support.

For additional information regarding the Printer Working Group visit: http://www.pwg.org

Contact information:

IFX Web Page: http://www.pwg.org/qualdocs
IFX Mailing List: ifx@pwg.org

To subscribe to the ipp mailing list, send the following email:
1) send it to majordomo@pwg.org
2) leave the subject line blank
3) put the following two lines in the message body:
subscribe ifx
end

Implementers of this specification are encouraged to join the IFX Mailing List in order to participate in any discussions of clarifications or review of registration proposals for additional names. Requests for additional media names, for inclusion in this specification, should be sent to the IFX Mailing list for consideration.

Contents

Deleted: 1 July 2003

Deleted: 35

Table of Tables

1 Introduction

This document specifies an application of PDF (Portable Document Format) that has two important properties: First, it is an "image"-based format, and proper rendering of the document is represented by (binary or color) images. Second, the format is suitable for incremental generation and thus it is a "streaming" format. The subset is called "PDF/is", for "PDF Image-Streamable".

PDF/is is formally a subset of PDF 1.4, and is intended to be fully compatible with software that reads PDF 1.4. There are "profiles" of PDF/is, which are distinguished primarily by the methods if image compression and/or techniques employed. The representations of image data employed are specified in the PDF 1.4 language reference [pdf], which in turn describes the PDF representation of image data specified by ITU-T recommendations for black-and-white facsimile ([t.4], [t.6]), ISO/IEG specifications for digital compression and coding of continuous-tone still images [jpeg], and lossy/lossless coding of bi-level images [jbig2].

PDF/is is intended to be useful within the IPPFAX protocol [ifx], which is used to provide a synchronous, reliable exchange of image documents between senders and receivers. For this reason, PDF/is also includes an optional security features for digital signaturing.

2 Terminology

This section defines terminology used throughout this document.

2.1 Conformance Terminology

Capitalized terms, such as **MUST, MUST NOT, REQUIRED, SHOULD, SHOULD NOT, MAY, NEED NOT, OPTIONAL,** and **PROHIBITED**, have special meaning relating to conformance as defined in RFC 2119 [rfc2119] and [rfc2911] section 12.1. If an implementation supports the extension defined in this document, then these terms apply; otherwise, they do not. These terms define conformance to *this document (and [rfc2911]) only*; they do not affect conformance to other documents, unless explicitly stated otherwise. To be more specific:

REQUIRED (REQ) - an adjective used to indicate that a conforming PDF/is Producer or Consumer's implementation MUST support the indicated operation, object, attribute, or attribute value. See [rfc2911] "Appendix A - Terminology for a definition of "support".

RECOMMENDED (REC) - an adjective used to indicate that a conforming PDF/is Producer or Consumer's implementation SHOULD support the indicated operation, object, attribute, or attribute value.

OPTIONAL (OPT) - an adjective used to indicate that a conforming PDF/is Producer or Consumer's implementation MAY support the indicated operation, object, attribute, or attribute value.

PROHIBITED (PROH) - an adjective used to indicate that a conforming PDF/is Producer or Consumer's implementation MUST NOT support the indicated operation, object, attribute, or attribute value.

AS SPECIFIED – is used to indicate that a conforming PDF/is Producer or Render implementation MUST, MAY, or MUST NOT support the indicated operation, object, attribute, or attribute value as is defined in the indicated specification.

OR – a conjunction that specifies a logical 'or', implying that a choice of one or more of the choices specified.

2.2 Other Terminology

The following terms are introduced and capitalized in order to indicate their specific meaning:

Implement – The specified feature is present in the Document.

Support – A Producer has the capability of Implementing the feature specified, or the Consumer has the capability of understanding and acting on the Implementation.

Document – The PDF/is-formatted electronic representation of a set of one or more pages that the Sender sends to the Receiver.

Consumer – This is the agent (software, hardware or some combination) that converts the Document into a displayed or printed form.

Producer -- This is the agent (software, hardware or some combination) that creates the Document.

Forward-Reference – In indirect object reference (See [pdf] Section 3.2.9) or a Resource Name (See Section 4.10) that refers to an object that appears later in the Document.

Cache – Consumer's storage, either memory, disk, or the like, to hold Document data as it's received from the Producer.

Page-Relative Objects – Objects that are indirectly referenced (See [pdf] Section 3.2.9) by either a 'Page' Dictionary or through a chain of object references that start with a reference from a 'Page' Dictionary.

Discarded – An adjective that describes a PDF object. An object is 'Discarded' when the Consumer no longer has access to the data within the object in question.

Object Size – The number of bytes required to represent an object in the Document. The size is calculated by subtracting the offset of the first byte of the line following the "endobj" of the object in question, from the offset of the first byte of the *object number* (See [pdf] Section 3.2.9).

Imaging Area – For the Producer, the Imaging Area of a page is the area specified by the Page Dictionary's 'MediaBox'. The Producer should use the actual area images from the source media for the 'MediaBox'. This would be the size of the input media for an edge-to-edge scan, for example. For the Consumer, the Imaging Area is an area on the output media that will contain all of the page's image content (the "inking" area). The Consumer usually uses the output media's printable area as the Imaging Area but may constrain it further to match the Producer's Imaging Area.

Scaled Page – When the Consumer's Imaging Area does not match the Producer's Imaging Area within 1/72 of an inch in either height OR width, the page is considered to be a Scaled Page.

Horizontal Scaling Factor – The Horizontal Scaling Factor is equal to the Consumer's Imaging Area width divided by the Producer's Imaging Area width, but MUST be 1.0 for a non-Scaled Page.

Vertical Scaling Factor – The Vertical Scaling Factor is equal to the Consumer's Imaging Area height divided by the Producer's Imaging Area height, but MUST be 1.0 for a non-Scaled Page.

Originator Identifier – An Image XObject that indicates information about the originator of the Document. See the protocol spec referencing this specification for details on what the 'Originator Identifier' MUST contain.

Nearest-Neighbor Interpolation – A two-dimensional interpolation of pixel values in which the amplitude of the interpolated sample is the amplitude of its nearest neighbor.

Bilinear Interpolation – A two-dimensional linear interpolation of pixel values based on the four pixels in a 2 x 2 pixel neighborhood.

Bicubic Interpolation – A two-dimensional cubic interpolation of pixel values based on the 16 pixels in a 4 x 4 pixel neighborhood.

3 PDF Document Requirements

The following table specifies the required (REQ), prohibited (PROH), and optionally (OPT) Supported PDF objects/filters for a Producer and Consumer to be considered compliant with this specification. Requirements for a specific object/filter to be considered Supported can be found in the 'PDF Object Requirements' section of this specification.

Table 3-1: PDF Object Requirements

PDF Object/Filter	Producer	Consumer	Reference
'ASCIIHexDecode' Filter	PROH	PROH	[pdf] Section (3.3.1)
'ASCII85Decode' Filter	PROH	PROH	[pdf] Section (3.3.2)
'LZWDecode' Filter	PROH	PROH	[pdf] Section (3.3.3)
'RunLengthDecode' Filter	PROH	PROH	[pdf] Section (3.3.4)
Incremental Updates	PROH	PROH	[pdf] Section (3.4.5)
Functions	PROH	PROH	[pdf] Section (3.9)
File specification	PROH	PROH	[pdf] Section (3.10)
Graphics State Parameter Dictionaries	PROH	PROH	[pdf] Section (4.3.4)
Path objects	PROH	PROH	[pdf] Section (4.4)
'DeviceGray' Color Space	PROH	PROH	[pdf] Section (4.5.3)
'DeviceRGB' Color Space	PROH	PROH	[pdf] Section (4.5.3)
'DeviceCMYK' Color Space	PROH	PROH	[pdf] Section (4.5.3)
Pattern Color Space	PROH	PROH	[pdf] Section (4.5.5)
Separation Color Space	PROH	PROH	[pdf] Section (4.5.5)
DeviceN Color Space	PROH	PROH	[pdf] Section (4.5.5)
Pattern Objects	PROH	PROH	[pdf] Section (4.6)
Inline Image Objects	PROH	PROH	[pdf] Section (4.8.6)
Form Xobjects	PROH	PROH	[pdf] Section (4.9)
Postscript Xobjects	PROH	PROH	[pdf] Section (4.10)
Font Objects	OPT	OPT	[pdf] Section (5)
Transparency	PROH	PROH	[pdf] Section (7)

Name Tree	PROH	PROH	[pdf] Section (3.8.4)
Number Tree	PROH	PROH	[pdf] Section (3.8.5)
'FlateDecode' Filter	OPT	REQ	[pdf] Section (3.3.3)
'CCITTFaxDecode' Filter	REQ	REQ	[pdf] Section (3.3.5)
File Header	REQ	REQ	[pdf] Section (3.4.1)
Cross-Reference Table	REQ	REQ	[pdf] Section (3.4.3)
File Trailer	REQ	REQ	[pdf] Section (3.4.4)
Document Catalog	REQ	REQ	[pdf] Section (3.6.1)
Page Tree Nodes	REQ	REQ	[pdf] Section (3.6.2)
Page Dictionary	REQ	REQ	[pdf] Section (3.6.2)
Content Streams	REQ	REQ	[pdf] Section (3.7.1)
Resource Dictionaries	REQ	REQ	[pdf] Section (3.7.2)
Image XObjects	REQ	REQ	[pdf] Section (4.7)
'JBIG2Decode' Filter	OPT	REQ	[pdf] Section (3.3.6)
'DCTDecode' Filter	OPT	REQ	[pdf] Section (3.3.7)
Encryption Dictionary	PROH	PROH	[pdf] Section (3.5)
'DeviceGray' Color Space	PROH	PROH	[pdf] pg. 182, See "ICCBased Color Space" section of this specification.
'DeviceRGB' Color Space	PROH	PROH	[pdf] pg. 184, See "ICCBased Color Space" section of this specification.
'Lab' Color Space	PROH	PROH	[pdf] pg. 187
'ICCBased' Color Space	REQ	OPT, See 'ICCBased Color Space' Section.	[pdf] pg. 189
'Indexed' Color Space	OPT	REQ	[pdf] pg. 199
Masked Images	OPT	REQ	[pdf] Section (4.8.5)
Interactive Form Dictionary and Annotation Field Dictionary and Signature Dictionary (Security Profile <DIG-SIG>)	OPT	OPT	[pdf] Section (8.6.1-3) [pdf-ppk] Section (2)
Cached Objects	REQ	REQ	Section 3.4
Banding	OPT	REQ	Section 3.3.11.3
Document Information Dictionary	OPT	OPT	[pdf] Section 9.2.1

3.1 File Layout (Informative)

Given that a Document is fully compliant with this specification, the Document will, nominally, have the following layout:

Table 3-2: File Layout

	Object
A	'PDF/is' Dictionary.
B	Page Dictionary for page 'n'
C	Content Stream 'a' for page 'n'
D	Image XObject 'x' for page 'n', stream 'a'
E	Color Space for image 'x' (cached), if not already loaded
F	Image Mask for image 'x', stream 'a', page 'n', if image is masked
G	[Repeat D-F for next Image 'x+1', stream 'a', page 'n', if present]
H	[Repeat C-G for next stream 'a+1' on page 'n', if present]

I	Content Stream Array for page 'n' (See Page Dictionary)
J	Resource Dictionary for page 'n'.
K	[Repeat B-J for next page 'n+1', if present]
L	Document Catalog
M	Page Tree Node(s)
N	Interactive Form Dictionary (If digitally signed)
O	Annotation Field Dictionary (If digitally signed)
P	Signature Dictionary (If digitally signed)
Q	Cross-Reference Table (See [pdf] Section 3.4.3)
R	File Trailer

4 PDF Object Requirements

The following sub-sections describe the object field values of the REQUIRED and OPTIONAL PDF objects in PDF/is. The numbers in '( )'s refer to section numbers in the PDF Specifications [pdf], unless otherwise noted. 'AS SPECIFIED' refers to the PDF Specification [pdf] unless otherwise noted.

All 'Required' and 'Optional' fields of a Document object (either specified here or referred to as 'Required' or 'Optional' in [pdf] or [pdf-ppk]) MUST be Supported if the object in question is to be considered 'Supported by the Consumer'. This rule does not apply if the definition of an object specifically states the requirements for the Consumer.

Support for all 'Required' fields of a Document object (either specified here or referred to as 'Required' in [pdf] or [pdf-ppk]) is REQUIRED if the object in question is to be considered 'Supported by the Producer'. Support for all 'Optional' fields of a Document object is OPTIONAL for the Producer. This rule does not apply if the definition of an object specifically states the requirements for the Producer.

4.1 'PDF/is' Dictionary

The 'PDF/is' Dictionary is a new Dictionary object that is REQUIRED for a PDF/is document.

The existence of this dictionary object is the one and only way to determine if the PDF in question is a PDF/is Document. The references in this object to items referred to in the Document Trailer are necessary to satisfy 'Producer Requirement' #6, see Section 4.1.

Table 4-1: PDF/is Dictionary

Field	Type	Specification
'Type'	Name	MUST have a value of '/Fis_PDFis'.
'Fis_Version'	Number	REQUIRED: A Real number of the format **MAJ_VER.MIN_VER**. (See below)
'Info'	Dictionary	MUST have same value as 'Info' field in the 'Document Trailer'. See [pdf] Table 3.12 for specification.
'ID'	Array	MUST have same value as 'ID' field in the 'Document Trailer'. See [pdf] Table 3.12 for specification.
'Fis_NextPage'	Dictionary	REQUIRED: MUST be an Indirect Object Reference to the first 'Page Dictionary'.
'Fis_DSig'	Dictionary	OPTIONAL: MUST be an Indirect Object Reference to the

Deleted: 1 July 2003

		'Signature Dictionary', if present.
'Fis_OrigID'	Dictionary	OPTIONAL: MUST be an Indirect Object Reference to the 'Originator Identifier' Image XObject, if present.
'Fis_Duplex'	Boolean	REQUIRED: MUST be 'false' unless the Document is known to be duplex and all odd numbered pages precede all even numbered pages (1, 3, 5,…, n*2 - 1, 2, 4, 6, …, n*2) – note that the last page (n*2) is optional since the Document may have an odd number of pages. See 'Page Ordering'.

See [pdf] Section 3.2.5 for definition of an 'Array Object'. See [pdf] Section 3.2.2 for definition of a 'Numeric Object'.

4.1.1 Fis_PDFis Key

4.1.1.1 MAJ_VER:

The 'major' version number of this PDF/is specification to which the Producer conforms to at the time the Document was created. The 'major' version of this specification is currently '1'.

4.1.1.2 MIN_VER:

The 'minor' version number of this PDF/is specification to which the Producer conforms to at the time the Document was created. The 'minor' version of this specification is currently '0'.

4.1.1.3 Example

An example of the PDF/is Dictionary for an encrypted, digitally signed, Document that needs a 4 Megabyte cache might look like this:

```
1 0 obj
<<
        /Type /Fis_PDFis
        /Fis_Version 1.0
        /Encrypt 2 0 R
        /Root 3 0 R
        /Info 4 0 R
        /ID [<8c41995c6e014675e850d36e6c2f6114><8c41995c6e014675e850d36e6c2f6114>]
        /Fis_NextPage 5 0 R
        /Fis_DSig 6 0 R
>>
endobj
```

4.2 PDF/is Format Identification

To refer to this version of the PDF/is specification from another specification, the string "PDF/is-1.0" should be used.

4.3 'CCITTFaxDecode' Filter

See [pdf] Section 3.3.5, [t.4], and [t.6]. Note that only 'Group 4' images are Supported by PDF/is, see 'K', below.

Table 4-2: CCITTFaxDecode Filter

Field	Specification
'K'	MUST have a value of -1.
'EndOfLine'	AS SPECIFIED
'EncodedByteAlign'	AS SPECIFIED
'Columns'	AS SPECIFIED
'Rows'	AS SPECIFIED
'EndOfBlock'	AS SPECIFIED
'BlackIs1'	AS SPECIFIED
'DamagedRowsBeforeError'	AS SPECIFIED

4.4 'JBIG2Decode' Filter

See [pdf] Section 3.3.6, [jbig2], and [t.89].

Table 4-3: JBIG2Decode Filter

Field	Specification
<All Details>	AS SPECIFIED, except as noted below.

- Consumers MUST support **Profile 1** (0x00000101 BASE), **Profile 2** (0x00000102 Upper Huffman), **Profile 3** (0x00000103 Lower Arithmetic) and **Profile 4** (0x00000104 Medium lossy/lossless arithmetic) as defined in [t.89]. Support for JBIG2 is OPTIONAL for the Producer. The Producer MUST NOT Implement any profile other than one of the four specified, above.
- All Consumers MUST support at least "Level 2" Memory (See [t.89], Table 1, Item 18).
- The Producer MUST adhere to the Function and Memory constraints as specified in [t.89].

4.5 'DCTDecode' Filter

See [pdf] Section 3.3.7, [ps-jpeg], [ps], and [jpeg].

PDF/is supports both the JPEG Baseline DCT and Extended sequential DCT compressed image formats.

Table 4-4: DCTDecode Filter

Field	Specification
<All Details>	AS SPECIFIED, except as noted below.

- Images MUST NOT be encoded using 'Progressive JPEG'.
- Images MUST have either 1 or 3 color components.
- All 3 component images (RGB, or YUV) MUST have their component data 'interleaved'. See [jpeg] Section 4.8.1.

- YUV encoding (See [pdf] pg. 60) is the RECOMMENDED encoding for image data. Rationale: Separation of luminance and chrominance information can facilitate greater image compression and simplifies the process of converting color image data to grayscale for Consumers that do not support color.

- The Consumer MUST adhere to the Memory requirements specified in Section 11 "RAM Requirements" of [ps-jpeg] for the Consumers Supported image resolution(s).

4.6 'FlateDecode' Filter

See [pdf] Section 3.3.3.

'Flate' encoding MUST NOT be used to compress image data. 'Flate' MAY only be used to compress non-image stream data, such as 'ICCBased Color Space' data, 'Indexed Color Space' data, and 'Content Stream' data.

See [pdf] Table 3.7:

Table 4-5: FlateDecode Filter

Field	Specification
<All Fields>	PROHIBITED.

4.7 File Trailer

See [pdf] Table 3.12.

Table 4-6: File Trailer

Field	Specification
'Size'	AS SPECIFIED
'Prev'	PROHIBITED
'Root'	AS SPECIFIED
'Encrypt'	PROHIBITED
'Info'	OPTIONAL.
'ID'	REQUIRED. MUST use a pseudo-random number in place of 'File Size' when generating this value. See [pdf] Section 9.3 for guidelines on how to generate this value. Rationale: Using a random number in place of file size is due to the requirements of using this field in generating the encryption key for the 'standard encryption' algorithm ([pdf] Step 5 of Algorithm 3.2, pg. 78): file size will not be known at the time this field is needed. Support for 'standard encryption' may be added to a future version of this specification.

4.8 Document Catalog

See [pdf] Table 3.16.

It should be noted that Page Attributes MUST NOT be Inherited (See [pdf] pg. 91) due to the nature of the ordering of the objects in this format. Rationale: Since the parent object (a Page Tree Node) of a Page Dictionary will not appear in the Document until after the page, streaming of the data for a page that has an inherited attribute would not be possible.

Table 4-7: Document Catalog

Field	Specification
'Type'	AS SPECIFIED
'Version'	AS SPECIFIED
'Pages'	AS SPECIFIED
'PageLabels'	PROHIBITED
'Names'	PROHIBITED.
'Dests'	PROHIBITED.
'ViewerPreferences'	OPTIONAL for both Producer and Consumer.
'PageLayout'	OPTIONAL for both Producer and Consumer.
'PageMode'	OPTIONAL for both Producer and Consumer.
'Outlines'	PROHIBITED.
'Threads'	PROHIBITED.
'OpenAction'	PROHIBITED.
'AA'	PROHIBITED.
'URI'	PROHIBITED.
'AcroForm'	REQ if <DIG-SIG>, PROH otherwise. MUST point to a 'Interactive Form Dictionary'
'Metadata'	AS SPECIFIED.
'StructTreeRoot'	PROHIBITED.
'MarkInfo'	AS SPECIFIED., See below.
'Lang'	PROHIBITED.
'SpiderInfo'	PROHIBITED.
'OutputIntents'	PROHIBITED.
'Fis_header	MUST be an indirect object reference to the 'PDF/is Dictionary'.

4.9 Page Tree Nodes

See [pdf] Table 3.17.

Table 4-8: Page Tree Nodes

Field	Specification
'Type'	AS SPECIFIED
'Parent'	AS SPECIFIED
'Kids'	AS SPECIFIED
'Count'	AS SPECIFIED
<All 'Page Dictionary' Fields, see [pdf] Table 3.18>	PROHIBITED

If the Producer of a Document knows that the Document is being generated in some non sequential order, this fact SHOULD be conveyed by reordering the 'Kids' objects from the order in which they appear in the Document. Rationale: If the Producing device were scanning the pages of a duplexed document by scanning the fronts of all pages first (as an example), reordering the

438 'Kids' objects in this way would allow a Consumer that has random access to the Document (i.e.
439 does not need to stream the data) the ability to display the pages in the proper order. If
440 reordering is to be accomplished, the Page Dictionary of the front and back of the same page
441 must have the same 'Parent' (Page Tree Node) entry in order to facilitate reorder, since all 'Kids'
442 of a particular Page Tree Node have sequential page numbers.

443

444 ## 4.10 Page Dictionary

445 See [pdf] Table 3.18.

446 **Table 4-9: Page Dictionary**

Field	Specification
'Type'	AS SPECIFIED
'Parent'	AS SPECIFIED
'LastModified'	AS SPECIFIED
'Resources'	MUST NOT be inherited, otherwise AS SPECIFIED.
'MediaBox'	MUST NOT be inherited, otherwise AS SPECIFIED.
'CropBox'	PROHIBITED: Same as 'MediaBox'.
'BleedBox'	PROHIBITED.
'TrimBox'	PROHIBITED.
'ArtBox'	PROHIBITED.
'BoxColorInfo'	PROHIBITED.
'Contents'	REQUIRED: MUST be an Indirect Object Reference to an Array Object that contains Indirect Object References to all Content Streams on the page. The Array Object MUST be placed immediately before the Resource Dictionary for the page.
'Rotate'	MUST NOT be inherited
'Group'	PROHIBITED.
'Thumb'	PROHIBITED.
'B'	PROHIBITED.
'Dur'	PROHIBITED.
'Trans'	PROHIBITED.
'Annots'	PROHIBITED.
'AA'	PROHIBITED.
'Metadata'	AS SPECIFIED.
'PieceInfo'	AS SPECIFIED.
'StructParents'	PROHIBITED.
'ID'	PROHIBITED.
'PZ'	OPTIONAL for both Producer and Consumer.
'SeparationInfo'	PROHIBITED.
'Fis_NextPage'	REQUIRED: An Indirect Object Reference to either: the next 'Page Dictionary'; or, if this is the last page in the Document, to the 'Document Catalog'.
'Fis_Duplex'	OPTIONAL: A 'boolean' object that defaults to 'false' and MUST be 'false' unless 'Fis_Duplex' in the 'PDF/is Dictionary' is 'true' and this is the first even numbered page in the Document.
'Fis_NextCS'	REQUIRED: MUST be an Indirect Object Reference to the first 'Content Stream' on the page.

447

4.10.1 Page Ordering

The Producer SHOULD order the pages in the Document sequentially from 1 to 'n'. For example, if the original document is duplex, the Producer SHOULD attempt to place the content from the back of page 1 (page 2) immediately after the content from page 1. This is preferable to placing content from all page fronts (odd number pages) followed by the content from all page backs (even numbered pages).

If the Producer chooses not to follow this page ordering guideline, the Producer MUST place all of the page fronts in the Document before all of the page backs – all odd numbered pages MUST precede all even numbered pages. In addition, the Producer MUST indicate this fact by specifying '/Fis_Duplex true' boolean object in the PDF/is Dictionary. The point at which the pages are flipped MUST be indicated by placing the '/Fis_Duplex true' boolean object in the Page Dictionary of the first even numbered page.

4.11 Content Streams

See [pdf] Table 3.4.

Table 4-10: Content Streams

Field	Specification
'Length'	REQUIRED: MUST not be an Indirect Object Reference.
'Filter'	PROHIBITED.
'DecodeParms'	PROHIBITED.
'F'	PROHIBITED.
'FFilter'	PROHIBITED.
'FDecodeParms'	PROHIBITED.
'Fis_NextCS'	REQUIRED: MUST be an Indirect Object Reference to the next Content Stream for the current page or the 'Resource Dictionary' if this is the last Content Stream on the page.

The dictionary mapping of Resource Names to indirect object numbers used in the Content Streams and Resource Dictionary MUST follow the following rule:

All Resource Names (See [pdf] Section 3.7.2) MUST have their indirect object ID's as the trailing part of the Resource Name. Resource Names MUST NOT have any digits (0-9) anywhere else in their name. Names MUST start with a letter. Consumers SHOULD use this convention to avoid having to cache the entire page in order to gain access to the Resource Dictionary at the end of the page data. For example, a page with two images that are overlapping and masked, might look like this:

```
3 0 obj %Page dictionary for page 1
<<
      /Type /Page
      /Resources 4 0 R
      /Contents 5 0 R
      …
>>
endobj

6 0 obj     %Content for page 1
<</Length 45>>
stream
```

```
            …
            /Im7 Do     % Image object at object number 7
            /Im8 Do     % Image object at object number 8
            /Fis_NextCS 4 0 R %Points to Res. Dict. – only one CS.
      endstream
      endobj

      7 0 R
      <<
            /Type /XObject
            /Colorspace /Cs9  % Color space at object number 9.
            …
      >>
      stream
      …
      endstream
      endobj

      10 0 R
      <<
            /Type /XObject
            /Mask 8 0 R
            /Colorspace /Cs7
            …
      >>
      stream
      …
      endstream
      endobj

      7 0 obj     %Color Space
      <</Length 3450>>
      stream
      …
      endstream
      endobj

      8 0 obj     %Mask for image object 10.
      …
      endobj

      5 0 obj
      [6 0 R]     %Array of Content Streams.
      endobj

      4 0 obj     %Resources for page 1
      <<
            /XObject << /Im9 9 0 R
                        /Im10 10 0 R >>
            /ColorSpace << /Cs7 7 0 R >>
      >>
      endobj
      //Page 2 would begin here…
```

Rationale: Since Indirect Object References from within Resource Dictionaries are prohibited (See [pdf] Section 3.7.2) we need a way to refer to these objects without requiring full buffering of a page. By requiring the objects to be written this way, the Consumer can process the Content Stream(s) and their associated Images and Color Spaces without requiring the Resource Dictionary. The Resource Dictionary must be written at the end of the page since it must refer to all objects that were used on the page.

See [pdf] Table 4.1:

Table 4-11: Content Stream Operators

Operators	Specification	Reference
q	AS SPECIFIED	[pdf] Table 4.7
Q	AS SPECIFIED	[pdf] Table 4.7
cm	MUST be [Sx 0 0 Sy Tx Ty], See Below	[pdf] Table 4.7
Do	AS SPECIFIED	[pdf] Table 4.34
DP	PROHIBITED except for 'Banding operator' and 'Cache operator', see below	[pdf] Table 9.8
BX	AS SPECIFIED	[pdf] Table 3.20
EX	AS SPECIFIED	[pdf] Table 3.20
BT	AS SPECIFIED	[pdf] Table 5.4
ET	AS SPECIFIED	[pdf] Table 5.4
'	AS SPECIFIED	[pdf] Table 5.6
"	AS SPECIFIED	[pdf] Table 5.4
T*	AS SPECIFIED	[pdf] Table 5.5
Tc	AS SPECIFIED	[pdf] Table 5.2
Td	AS SPECIFIED	[pdf] Table 5.5
TD	AS SPECIFIED	[pdf] Table 5.5
Tf	AS SPECIFIED, also see Font Objects	[pdf] Table 5.2
Tj	AS SPECIFIED	[pdf] Table 5.6
TL	AS SPECIFIED	[pdf] Table 5.2
Tm	AS SPECIFIED	[pdf] Table 5.5
Tr	REQUIRED, and MUST be '3'	[pdf] Table 5.2
Ts	AS SPECIFIED	[pdf] Table 5.2
Tw	AS SPECIFIED	[pdf] Table 5.2
Tz	AS SPECIFIED	[pdf] Table 5.2
<All other Operators>	PROHIBITED	[pdf] Table A.1

Support for text operators (all operators beginning with the letter 'T', as well as the BT, ET, ', and " operators) are OPTIONAL for both the Producer and the Consumer. If text operators are found in a Document, the Consumer MAY ignore them as they do not affect the rendering of the page content since all text MUST be 'invisible' (Text Mode (Tr) == 3).

4.11.1 'cm' Operator:

See [pdf] Table 4.7 for definition of 'cm' operator. Note that all coordinates in PDF/is are in the 'default user space' (See [pdf] pg. 138).

Given:

Wi = Width (X-direction) of the Image in inches.

Hi = Height (Y-direction) of the Image in inches.

Xi = Horizontal translation, in inches, from the left edge of the page to the left edge of the image.

Yi = Vertical translation, in inches, from the bottom edge of the page to the bottom of the image.

The Producer MUST ensure that the following is true:

Sx = Wi * 72

Sy = Hi * 72

Tx = Xi * 72

Ty = Yi * 72

4.11.2 'Do' Operator:

See [pdf] Table 4.34 for definition of 'Do' operator.

Image Resolution Calculations

Given:

Img = The 'Image XObject' associated with the 'Do' operator.

Cm = The current 'cm' operation in effect for 'Img'.

Wp = 'Width' field of 'Img'.

Hp = 'Height' field of 'Img'.

Sx = 'Sx' value of 'Cm'.

Sy = 'Sy' value of 'Cm'.

The following must be assumed by the Producer and the Consumer:

(Wp * 72 / Sx) = The resolution, in the X-direction, of 'Img', in dots per inch.

(Hp * 72 / Sy) = The resolution, in the Y-direction, of 'Img', in dots per inch.

4.11.3 'DP' Operators:

See [pdf] Table 9.8 for a definition of the 'DP' Operator.

Only the 'Marked Content' flags 'Banding Operator' and the 'Cache operator' are permitted in PDF/is, all other flags are PROHIBTED.

4.11.3.1 'Banding' Operator:

Banding facilitates the creation of a complex series of images on a PDF/is page to a Consumer that may be memory constrained and unable to otherwise display the page. If the Producer of the Document is able to determine that the current page's image layering (or "masking") will violate the cache memory constraints of the Consumer; the Consumer MUST break up the current page into non-overlapping regions to be displayed ('Banding') or free up resources using the 'Cache Operator' (see below). Banding is specified in one of the content streams of the page.

All images or masks in the content stream in a particular 'Band' do not overlay, and are not overlaid by, any images or masks in any other 'Band'.

To indicate that a new 'Band' is beginning, the content stream MUST contain the following operator syntax, exactly as shown:

/Fis_band<</Fis_band [Y]>> DP

Where:

Y: A 'Real Numeric Object' (See [pdf] Section 3.2.2) of the minimum Y-coordinate value that this band will contain.

And:
All coordinate values are in the 'default user space' (See [pdf] pg. 138) coordinate system (0,0 is lower left), at 72 units per inch, relative to the Page Dictionary's 'MediaBox'.

- Bands may only progress from top to bottom (highest to lowest Y coordinate).
- The last Band on the page MUST not have a Banding operator since the close of the Content Stream will indicate that the last band is to be rendered.
- The extent of an image within a particular Band MUST meet the following requirements:
  - Its top edge MUST have a y-coordinate value less than the **Y** value of the previous Band.
  - Its bottom edge MUST have a y-coordinate greater than, or equal to the **Y** value of the current Band, or '0' if this is the last band.

See the following examples to help illustrate this feature.

For the examples, below:
N: [Y]
Where 'N' is the order in which the band appears in the Content Stream.
'Y' is the 'Y' value of the Band operator.

Example #1: an 8.5" X 11" page (612x792 units), divided into 3 equal sized Bands:

1: [528]
2: [264]
3: (No operator)

Example #2: and 11" X 17" page (792x1224 units), divided into 4 "bands":

1: [918]
2: [612]
3: [306]
4: (No operator)

A 'Band Operator' MAY occur in any Content Stream for that page. If the page has more than one Content Stream it MUST be considered as described in [pdf] page 89, under 'Contents'.

To illustrate what a 'Banded' content stream might look like; here is the content stream for Example #2, above:

```
stream
q
792 0 0 306 0 1224 cm   % region of first 'band'.  792 units
wide, 306 units high,
/Im1 Do                 % Display image in first band.
/Fis_band <</Fis_band [918]>> DP    % 'Band Operator'
```

```
Q
q
792 0 0 306 0 918 cm
/Im2 Do                 % Display image in second band.
/Fis_band <</Fis_band [612]>> DP
Q
q
792 0 0 306 0 612 cm
/Im3 Do                 % Display image in third band.
/Fis_band <</Fis_band [306]>> DP
Q
q
792 0 0 306 0 306 cm
/Im4 Do                 % Display image in last band.
endstream
```

4.11.3.2 'Cache' Operator:

The 'Cache Operator' allows the Producer of the Document to specify that certain 'cached' objects (See 'Cached Objects' section in this specification) may be released from the cache at a certain point in the content stream. See 'Cache Release' section in this document for use of this operation. This operation would allow a Consumer to Discard specified objects to free resources for image operations. This operator has the following syntax:

```
/Fis_cache <</Fis_cache [OBJECTS]>> DP
```

Where 'OBJECTS' is an array of object ID references. For example:

```
/Fis_cache <</Fis_cache [23 0 R 34 0 R]>> DP
```

…will release objects 23 and 34 from the cache.

4.12 Resource Dictionaries

See [pdf] Table 3.21.

The Resource Dictionary MUST reference all Image XObjects and ColorSpaces that are used on the current page. The position of the image objects, their masks, and color spaces with respect to each other is defined in the Image XObject section of this specification.

The 'Resource Dictionary' MUST be the last object for any given page. This is an indicator to the Consumer that the current page is complete.

Table 4-12: Resource Dictionaries

Field	Specification
'ExtGState'	PROHIBITED.
'ColorSpace'	PROHIBITED.
'Pattern'	PROHIBITED.
'Shading'	PROHIBITED.
'XObject'	AS SPECIFIED.
'Font'	AS SPECIFIED.
'ProcSet'	PROHIBITED.
'Properties'	PROHIBITED.

4.13 ICCBased Color Space

See [pdf] Table 4.16 & Table 3.4.

Table 4-13: ICCBased Color Space

Field	Specification
'N'	MUST have a value of '3'.
'Alternate'	PROHIBITED, Implies '/DeviceRGB' (See [pdf]).
'Range'	AS SPECIFIED.
'Metadata'	AS SPECIFIED.
'Length'	REQUIRED. MUST NOT be an indirect object reference.
'Filter'	PROHIBITED.
'DecodeParms'	PROHIBITED.
'F'	PROHIBITED.
'FFilter'	PROHIBITED.
'FDecodeParms'	PROHIBITED.

The following rules MUST be adhered to:

- All color image data MUST be 'sRGB' color data (See [srgb]). Color images MUST use the 'sRGB' standard ICC profile [srgb-icc].
- The [srgb-icc] profile MUST be Implemented in the Document, unmodified.
- The profile MUST be Implemented after its first reference (See Producer Conformance Requirement #6) and SHOULD be cached (See 'Cached Objects') for further references.

Since the color image data meets the 'sRGB' specification, the Consumer has the following two options:

1 Tune the output device to use 'sRGB' image data. This would allow the Consumer to avoid having to implement a full ICC profile engine. The image data would be used directly which could greatly simplify the image data processing.

2 Support ICC profiles. In this case, the Consumer does not need to know that the image data conforms to 'sRGB'; instead, the Consumer can process the data using an entirely ICC based color management approach (See [icc]). This method would be the choice for the Consumer that supports the full PDF specification [pdf].

4.14 Indexed Color Space

See [pdf] Page 199.

An Indexed color space MAY be used for grayscale or color images, as necessary.

An Indexed Color Space object MUST take the following form:

[/Indexed *base hival lookup*]

Where:

'base' MUST be an array of the form:

[/ICCBased *X*]

Where 'X' is an indirect object reference to an ICCBased 'sRGB' color space (See ICCBased Color Space).

'hival' MUST be as defined on page 200 in [pdf].

'lookup' MUST be as defined on page 200 in [pdf] but MUST be a stream.

Example:

```
10 0 obj
[/Indexed [/ICCBased 12 0 R] 255 11 0 R]]
endobj

11 0 obj
<</Length 768>>
stream
….%256 color lookup table values in R-G-B order…
endstream
endobj

12 0 obj
%ICCBased 'sRGB' color space
…
```

4.15 Image XObjects

See [pdf] Table 4.35 & Table 3.4 for description of the following table.

Table 4-14: Image XObjects

Field	Specification
'Type'	MUST be 'XObject'
'Subtype'	MUST be 'Image'
'Width'	AS SPECIFIED
'Height'	AS SPECIFIED
'ColorSpace'	AS SPECIFIED. Only 'ICCBased' or 'Indexed' color spaces are permitted.
'BitsPerComponent'	AS SPECIFIED
'Intent'	REQUIRED. 'Perceptual' is RECOMMENDED.
'ImageMask'	AS SPECIFIED
'Mask'	AS SPECIFIED, see below.
'SMask'	PROHIBITED.
'Decode'	AS SPECIFIED.
'Interpolate'	AS SPECIFIED. 'False' implies "Nearest-Neighbor Interpolation". 'True' implies 'Bilinear Interpolation' or 'Bicubic Interpolation' at the discretion of the Consumer. The actual method by which these are implemented is not specified.
'Alternates'	PROHIBITED.
'Name'	PROHIBITED.
'StructParent'	PROHIBITED.
'ID'	PROHIBITED.
'OPI'	PROHIBITED.
'Metadata'	AS SPECIFIED.
'Length'	REQUIRED: MAY be an indirect object reference to a numeric object that

	MUST be the next object in the Document, See below.
'Filter'	REQUIRED: MUST be one of: 'DCTDecode', 'CCITTFaxDecode', or 'JBIG2Decode'. No other filters are allowed.
'DecodeParms'	AS SPECIFIED.
'F'	PROHIBITED.
'FFilter'	PROHIBITED.
'FDecodeParms'	PROHIBITED.

- An 'ImageMask', if indicated in an Image XObject, MUST appear in the Document before the Image XObject that references it.
- All image data, regardless of compress method (Filter), MUST be ordered as specified in Section 4.8.3 and in Figure 4.26 of [pdf], contrary to the 'Note' at the bottom of page 265 of [pdf].
- Grayscale images MUST use an Indexed Color Space.
- If the 'Length' specifier for a stream is an indirect object reference to a numeric object, the Producer MUST place the following comment on the line after the 'endstream' keyword:
    - %ID['ID' field value from 'PDF/is Dictionary']

  Using Section 4.1.1.3 as an example, we would have:

  endstream
  %ID[<8c41995c6e014675e850d36e6c2f6114><8c41995c6e014675e850d36e6c2f6114>]

  Rationale: By placing this 'ID' at the end of the stream object a Consumer does not have to understand the format of the stream in order to find its end. The Consumer can simply search for the 'ID' string to determine where the stream ends. This is mainly useful when the Consumer is reading a newer version of the PDF/is document format that it does not understand.

4.16 Masked Images

See [pdf] Section 4.8.5.

Table 4-15: Masked Images

Field	Specification
<All Fields>	AS SPECIFIED

4.17 Interactive Form Dictionary

See [pdf] Table 8.47.

Table 4-16: Interactive Form Dictionary

Field	Specification
'Fields'	MUST be an Array of indirect object reference(s) to 'Annotation Field Dictionary'(s).
'NeedAppearances'	PROHIBITED
'SigFlags'	MUST be '3'

'CO'	PROHIBITED
'DR'	PROHIBITED
'DA'	PROHIBITED
'Q'	PROHIBITED

4.18 Font Objects

'Font Objects' (See [pdf] Section 5.4) include both 'Font Dictionaries' ([pdf] Table 5.8) and 'Font Descriptors' ([pdf] Table 5.18).

Fonts can be used in PDF/is Documents only for text searching and extraction capabilities. All text MUST be invisible (See 'Tr' in Content Streams). As such, support for Font Objects is OPTIONAL for both the Producer and the Consumer. Since text is invisible, the Consumer need not Support Text Operators (in Content Streams) or Font Objects as they do not affect the rendered output.

Font Objects, if present, MUST follow the following rules:

- Embedded font programs ([pdf] Section 5.8) are PROHIBITED.
- All font 'SubTypes' ([pdf] Table 5.7) except 'TrueType' ([pdf] Section 5.5.2) and 'Type1' ([pdf] Section 5.5.1) are PROHIBITED.
- 'Font Dictionaries' MUST be implemented AS SPECIFIED in [pdf].
- 'Font Descriptors' MUST be Implemented AS SPECIFIED in [pdf].

4.19 Annotation Field Dictionary

See [pdf] Tables 8.10 & 8.49. This dictionary consists of entries from both a 'Annotation Dictionary (Table 8.10) and a 'Field Dictionary' (Table 8.49).

Only Digital Signature Annotations are allowed in PDF/is.

Table 4-17: Annotation Field Dictionary

Field	Specification
'Type'	MUST be 'Annot'
'Subtype'	MUST be 'Widget'
'Contents'	PROHIBITED.
'P'	PROHIBITED.
'Rect'	MUST be '[0 0 0 0]'
'NM'	PROHIBITED.
'F'	PROHIBITED.
'BS'	PROHIBITED.
'Border'	PROHIBITED.
'AP'	PROHIBITED.
'AS'	PROHIBITED.
'C'	PROHIBITED.
'CA'	PROHIBITED.
'T'	PROHIBITED.
'Popup'	PROHIBITED.
'A'	PROHIBITED.

'AA'	PROHIBITED.
'StructParent'	PROHIBITED.
'FT'	MUST be 'Sig'
'Parent'	PROHIBITED.
'Kids'	PROHIBTED.
'T'	AS SPECIFIED.
'TU'	AS SPECIFIED.
'TM'	PROHIBITED.
'Ff'	MUST be '1'.
'V'	MUST be an indirect object reference to a 'Signature Dictionary'.
'DV'	PROHIBITED.
'AA'	PROHIBITED.

4.20 Signature Dictionary

See [pdf] Table 8.60 and [pdf-ppk] Table 2.

The Digital Signature format MUST only be in the 'Raw Format', see [pdf-ppk] Section 2.2.

Table 4-18: Signature Dictionary

Field	Specification
'Type'	MUST be 'Sig'
'Filter'	AS SPECIFIED.
'SubFilter'	MUST be 'adbe.x509.rsa_sha1'
'Name'	AS SPECIFIED.
'Reason'	AS SPECIFIED.
'Location'	AS SPECIFIED.
'M'	AS SPECIFIED.
'ByteRange'	PROHIBITED (Implies all bytes in the Document with the exclusion of the bytes represented by the value of the 'Cert' field. See [pdf] for this field)
'Contents'	AS SPECIFIED.
'Cert'	AS SPECIFIED.
'R'	AS SPECIFIED.
'V'	AS SPECIFIED.
'ADBE_Build'	AS SPECIFIED.
'ADBE_AuthType'	AS SPECIFIED.
'ADBE_PwdTime'	AS SPECIFIED.

5 Object Lifetime

Some Consumer's may be limited in the amount of storage they may have to cache the Document as it's received from the Producer. This storage limitation may prohibit the Consumer from holding the entire Document before beginning to render the first page. To facilitate this storage constraint, PDF/is has a mechanism of "object lifetime". This mechanism defines how long an object must be held in storage before it is no longer needed.

If a Document can be fully maintained in the Consumer's storage, i.e. the Consumer is a PC or some other device with large quantities of storage; the Document's Cross-Reference table should

be used to access objects as they are needed. In this case, the Consumer should follow the parsing model as spelled out in the PDF Reference [pdf].

If a Document cannot be fully maintained within the Consumers storage or if it is uncertain if it will be able to do so, the Document MUST be linearly parsed and the following parsing rules MUST be adhered to:

- Documents MUST be parsed in order, from beginning to end.
- All Consumer's MUST have the ability to cache at least 4 Megabytes (4,194,304 bytes) of PDF/is Document data. This memory is in addition to any memory required for JBIG2 image processing (2 Megabytes, See 'JBIG2Decode' Section) and for raster image buffers on the Consuming device.

At the end of generation of each Dictionary Object (See [pdf] Section 3.2.6), the Producer MUST ensure that 4 Megabyte cache memory limit will not been exceeded when the Consumer reads the Document. If the Producer exceeds the limit as calculated using the formula shown below, the Document is Invalid. If the limit will be exceeded, the Producer MUST either reorganize the current page by using either "Banding", freeing up some "cached" objects, reducing the use of masked images (or lowering their resolution), or by using some other process in order to avoid breaking the cache buffer limit.

Calculation of the current cache buffer size MUST follow the following formula:

1) The current total Document size (in bytes) that has been created up to the point at which this calculation is being made.
2) Minus the 'Object Size' of all released 'Cached' objects (See "Cached Objects" Section of this specification), up to that point.
3) Minus the 'Object Size' of all non-cached 'Page-Relative Objects' for previous pages, not already accounted for by #2.
4) Minus the 'Object Size' of all non-cached 'Image XObjects' data for any previous 'Bands' on the current page; if the page is "Banded".
5) Minus the 'Object Size' of the last 'Image XObject' in the current 'Band', if the page is "Banded".
6) Minus the 'Object Size' of the 'Image XObject' for the current page, if the page is not "Banded".

   Rationale: The last two items assume that the Consumer will process image data as it is received and will not need to cache these objects before rendering.

6 Cached Objects

If a 'Page-Relative' object MAY be used on more than one page or in more than one 'Band', it will be necessary to specify the object as 'Cached'. This will allow an object to be used throughout the Document that otherwise would be discarded. This caching mechanism only applies to 'Page-Relative' 'Dictionary Objects'; see [pdf] Section 3.2.6.

An object that is held in the Consumers cache by the 'Cache Hold' mechanism MUST be maintained in the cache until one of the following conditions is met:

- The 'Cache Operator' is invoked on this object in a page's Content Stream.
- The 'Document Catalog' is reached.

Deleted: 1 July 2003

To specify that a particular object should be 'cached', add the following Name Object (See [pdf] Section 3.2.4) to the Dictionary Object (See [pdf] Section 3.2.6) to be cached:

/Fis_Cache

7 Conformance Requirements

This section specifies the conformance requirements for Consumers and Producers.

7.1 Producer conformance requirements

In order to conform to this specification, a Document Producer:

1. MUST specify the version of PDF (See [pdf] Section 3.4.1) as being 'PDF 1.4'.
2. MUST place the 'PDF/is Dictionary' as the first object in the PDF.
3. MUST NOT include any private 'PDF Name Registry' values/objects (See [pdf] – Appendix E) that affect printed output.
4. MUST place the objects: 'Interactive Form Dictionary', 'Annotation Field Dictionary' and 'Digital Signature' objects as the last three objects (in that order) in the Document, if the Document is Digitally Signed. Note that in a situation where the Consumer cannot cache the entire document before rendering, the detection of a valid or invalid Digital Signature will only occur after rendering of the entire Document.
5. MUST ensure that there is at least one Forward-Reference to each object. The only object that does not have to follow this rule is the 'PDF/is Dictionary'. Rationale: This will aid the Consumer with identifying objects as they are encountered in the data stream.
6. MUST ensure that all objects appear in the PDF AFTER the object in which they are first referenced (Satisfied by Requirement 6) and BEFORE the next 'Page Dictionary' unless the object is a Cached Object (See Section 3.4).
7. MUST ensure that all object identifiers ([pdf] Section 3.2.9) start at the beginning of a line.
8. MUST ensure that all 'endobj' keywords ([pdf] Section 3.2.9) start at the beginning of a line.
9. MUST NOT Linearize the Document. See [pdf] Appendix F.
10. MUST NOT Incrementally Update the Document. See [pdf] Section 3.4.5.
11. MUST only encoded images with resolutions of at least 300 but not more than 1200 dots per inch (dpi). It is RECOMMENDED that the Producer place images in the Document in the images original resolution, i.e. not scaled.
12. MAY include an 'Originator Identifier' image that MUST, if present, be displayed on, at least, the first page. The image MUST be referenced by the 'Fis_OrigID' field in the 'PDF/is Dictionary' and MUST be 'cached' if displayed on more than the first page.
13. MUST end all text lines with a PDF Reference specified 'EOL Marker' (See [pdf] pg. 26).

Deleted: single line-feed (0x0A) character

Deleted: MUST NOT use a single carriage-return (0x0D) nor a carriage-return plus line-feed combination (0x0D, 0x0A) to signify the end of a line.

Deleted: 1 July 2003

14. MUST not use multiple, sequential 'EOL Markers' (See [pdf] pg. 26), i.e. there should be no blank lines in the Document.

15. MUST only use either a space or a horizontal tab character as white space ([pdf] Table 3.1).

16. MUST keep white-spaces to a single instance. Runs of multiple white-space characters are PROHIBITED.

17. MUST place the following five characters as the second line in the Document: %âãÏÓ (Hex values 0x25, 0xE2, 0xE3, 0xCF, 0xD3)

18. MUST separate the 'xfer' keyword from the cross reference subsection header by a single EOL Marker (See [pdf] Section 3.4.3).

19. MUST NOT place any data following the '%%EOF' at the end of the Document.

20. MUST NOT place any data between the end of one Dictionary object and the beginning of the next Dictionary object.

21. MUST place an 'EOL Marker' after all 'stream' keywords.

22. MUST place an 'EOL Marker' before all 'endstream' keywords.

23. MUST place an 'EOL Marker' after all 'obj' keywords.

24. MUST place an 'EOL Marker' after all 'endobj' keywords.

25. MUST place all *object numbers*, *generation numbers*, and 'obj' keywords (See [pdf] Section 3.2.9) together on a single line and the individual items are each to be separated by a single white space character.

7.2 Consumer conformance requirements

In order to conform to this specification, a Document Consumer:

1. MUST Support all of the REQUIRED objects.

2. MUST Interpolate images up or down in resolution, as required, to properly match the Document's image resolution(s) to the Consumer's device capabilities.

3. MUST abide by the "Object Lifetime" rules in Section 3.4 if unable to Cache the entire Document.

4. MUST terminate processing of the Document if it is detected that the Document has been incrementally updated (See [pdf] Section 3.4.5) as these Documents are PROHIBITED.

5. MUST have a Horizontal Scaling Factor that is within 0.3% of the Vertical Scaling Factor for any particular page.

6. MUST have all Vertical and Horizontal Scaling Factors within the range of 0.9 and 1.1, inclusive for all pages.

7. MAY display the Originator Identifier where specified in a page's Content Stream.

8. MUST attempt to recover from an invalid Document. Any Document that does not conform to this specification is considered to be 'Invalid'. If a formatting error is encountered in a Document, the Consumer MUST attempt to recover from the error by following the rules shown below.

    a. If the error was encountered in a stream, the Consumer MUST skip to the end of the stream ignoring all remaining data in the stream.

    b. If the error was encountered in an object outside of a stream, the Consumer SHOULD skip to the end of the current object, if possible. If not possible, the Consumer MUST skip to the next Page Object.

    It should be noted that skipping objects in this way will cause the current page to be invalid. The details of handling invalid pages is outside the scope of this specification. In addition, If some of the skipped objects were 'Cached' additional pages may also be invalid.

8 Issues

- None currently.

9 Sample PDF/is Document

The 'source' of the sample document in this section can be viewed with most text editors ('Wordpad' is a good choice) but should only be modified with a binary editor, as the stream data contained therein is not compatible with text editors. Comments on the format of the documents are contained within the documents themselves.

This sample is an one page document. The page contains a 'CCITTFaxDecode' masked, 'DCTDecode' color foreground image with a 'DCTDecode' gray scale background image.
ftp://pwg.org/pub/pwg/QUALDOCS/SamplePDFax/sample2.pdf

10 Normative References

[pdf]
Adobe Systems, "PDF Reference, third edition, Adobe Portable Document Format Version 1.4", Addison-Wesley, December 2001, http://partners.adobe.com/asn/developer/acrosdk/docs/filefmtspecs/PDFReference.pdf. Also see errata: http://partners.adobe.com/asn/developer/acrosdk/docs/PDF14errata.txt.

[pdf-ppk]
Pravetz, J., "PDF Public-Key Digital Signature and Encryption Specification", Version 3.2, Adobe Systems, September 2001, http://partners.adobe.com/asn/developer/pdfs/tn/ppk_pdfspec.pdf

[ps-jpeg]
Adobe Systems Incorporated, "Supporting the DCT Filters in PostScript Level 2", November 1992, http://partners.adobe.com/asn/developer/pdfs/tn/5116.DCT_Filter.pdf

[ps]
Adobe Systems Incorporated, “PostScript Language Reference third edition”, Addiseon-Wesley, 1999, http://partners.adobe.com/asn/developer/pdfs/tn/PLRM.pdf. Also see errata: http://partners.adobe.com/asn/developer/pdfs/tn/PSerrata.txt.

[ifx]
Moore, Songer, Hastings, Seeler “IPPFAX/1.0 Protocol” PWG Proposed Standard, (Work in Progress), ftp://pwg.org/pub/pwg/QUALDOCS/pwg-ifx-ippfax-latest.pdf

[ifx-req]
Moore, P., “IPP Fax transport requirements”, October 16, 2000, ftp://pwg.org/pub/pwg/QUALDOCS/requirements/ifx-transport-requirements-01.pdf

[t.4]
ITU-T Recommendation T.4, “Standardization of group 3 facsimile apparatus for document transmission”, October 1997

[t.6]
ITU-T Recommendation T.6, “Facsimile coding schemes and coding control functions for group 4 facsimile apparatus”, November 1988

[t.89]
ITU-T Recommendation T.89, “Application profiles for Recommendation T.88 – Lossy/lossless coding of bi-level images (JBIG2) for facsimile”, September 2001

[rfc2119]
Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, September 2000, ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc2911.txt.pdf.

[rfc2911]
Hastings, Herriot, deBry, Isaacson, Powell, “Internet Printing Protocol/1.1: Model and Semantics”, September 2000, ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc2911.txt.pdf.

[jpeg]
JTC 1/SC 29, “Information technology – Digital compression and coding of continuous-tone images: Requirements and guidelines”, ISO/IEC 10918-1:1994, 1994.

[jbig2]
JTC 1/SC 29, “Information technology – Lossy/lossless coding of bi-level images”, ISO/IEC 14492:2001, December 2001.

[icc]
International Color Consortium (ICC), ICC.1:1998-09, “File Format for Color Profiles”, 1998. http://www.color.org/ICC-1_1998-09.PDF

[icc-a]
International Color Consortium (ICC), ICC.1A:1999-04, “Addendum 2 to Spec. ICC.1:1998-09”, 1999. http://www.color.org/ICC-1A_1999-04.PDF

[srgb]
International Electrotechnical Commission (IEC), IEC/3WD 61966-2.1, “Colour Measurement and Management in Multimedia Systems and Equipment, Part 2.1: Default RGB Colour Space—sRGB”, 1999.

Deleted: 1 July 2003

[srgb-icc]
sRGB ICC Color Profile: “sRGB Color Space Profile.icm”.
http://www.srgb.com/usingsrgb.html

11 Informative References

[rfc2542]
Masinter , "Terminology and Goals for Internet Fax", RFC2542, March 1999, ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc2542.txt.pdf.

[ifx-goals]
Klyne, Shockey, "Additional Goals for Quality Document Transfer", October 1999, ftp://ftp.pwg.org/pub/pwg/QUALDOCS/Internet-Drafts/draft-klyne-qualdoc-goals-02.txt.

[pdf-a]
PDF-Archive Committee, “Document Management – Long-term electronic preservation – Use of PDF (PDF/A)”, May 2003, http://www.aiim.org/standards.asp?ID=25013.

[process]
“PWG Policy: Definition of the Standards Development Process”, April 2003, ftp://ftp.pwg.org/pub/pwg/general/process/pwg-process20-20030414.pdf

12 Revision History (to be removed when standard is approved)

Date	Author	Notes
10/9/02	Rick Seeler, Adobe Systems	Version 0.01 (never released)
10/23/02	Rick Seeler, Adobe Systems	Version 0.02 ftp://pwg.org/pub/pwg/QUALDOCS/pwg-ifx-pdfax-P02-021023-rev.pdf
11/19/02	Rick Seeler, Adobe Systems	Version 0.03 ftp://pwg.org/pub/pwg/QUALDOCS/pwg-ifx-pdfis-P03-021110-rev.pdf
11/22/02	Rick Seeler, Adobe Systems	Version 0.04 ftp://pwg.org/pub/pwg/QUALDOCS/pwg-ifx-pdfis-P04-021122-rev.pdf
12/19/02	Rick Seeler, Adobe Systems	Version 0.05 ftp://pwg.org/pub/pwg/QUALDOCS/pwg-ifx-pdfis-P05-021219-rev.pdf
2/19/03	Rick Seeler, Adobe Systems	Version 0.06 ftp://pwg.org/pub/pwg/QUALDOCS/pwg-ifx-pdfis-P06-030219-rev.pdf
3/14/03	Rick Seeler, Adobe Systems	Version 0.50 ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030314-rev.pdf
3/24/03	Rick Seeler, Adobe Systems	Version 0.60 ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030324-rev.pdf
5/6/03	Rick Seeler, Adobe Systems	Maturity: Prototype ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030506-rev.pdf
6/30/03	Rick Seeler, Adobe Systems	Maturity: Prototype

Deleted: Stable

Deleted: 1 July 2003

		ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030630-rev.pdf
8/5/03	Rick Seeler, Adobe Systems	Maturity: Prototype ftp://pwg.org/pub/pwg/QUALDOCS/wd-pdfis10-20030805-rev.pdf

13 Contributors

Rick Seeler - Adobe Systems mailto:rseeler@adobe.com
John Pulera - Minolta mailto:jpulera@minolta-mil.com
Gail Songer - Peerless mailto:gsonger@peerless.com
Tom Hastings - Xerox mailto:hastings@cp10.es.xerox.com
Rob Buckley - Xerox mailto:rbuckley@crt.xerox.com
Lloyd McIntyre mailto:lloyd10328@pacbell.net
Ira McDonald - High North mailto:imcdonald@sharplabs.com

Formatted: French (France)
Formatted: French (France)
Formatted: French (France)
Field Code Changed

14 Acknowledgments

Kari Poysa - Xerox mailto:Kari.Poysa@usa.xerox.com
Jerry Thrasher - Lexmark mailto:thrasher@lexmark.com
Don Wright - Lexmark mailto:don@lexmark.com
Martin Bailey - Global Graphics mailto:martin.bailey@globalgraphics.com

15 Author's Address

Rick Seeler
Adobe Systems Incorporated
321 Park Ave., E13
San Jose, CA 95110
Phone: 1+408 536-4393
Fax: 1+408 537-8077
e-mail: mailto:rseeler@adobe.com

16 Appendix A – Intellectual Property

In addition to this section, see the 'Intellectual Property' or 'Patent' sections in the specifications refered to by the Normative References in this specification for additional Intellectual Property related issues.

16.1 Patents – Unknown Status

The following patents have been brought forward as possibly relevant intellectual property pertaining to implementations of PDF/is. No formal statement has been made by the patent holder(s) as to the relevance of these patents with respect to implementations of PDF/is.

Patents listed here meet all of the following three criteria:

1) The patent has been identified by someone who is familiar with the technical fields relevant to this Specification, and who believes use of the invention covered by the patent may be infringed upon by a particular implementation of this Specification.

Deleted: 1 July 2003

2) The patent has not been identified as being essential to PDF/is: the patent will not necessarily be infringed upon by an implementation of PDF/is but some implementations may do so.

Deleted: non-

3) The patent holder is not currently willing to make the intellectual property freely available as defined in Item 1 under section 9.3 of the PWG Process Document [process].

Patents:

1) US Patent, RE35657, Xerox, Buckley et. al.: Means for combining data of different frequencies for a raster output device., Nov. 11, 1997.

Formatted: Bullets and Numbering

2) US Patent 5778092, Xerox, MacLeod et. al.: Method and apparatus for compressing color or gray scale documents., Dec. 20, 1996.

16.2 Patents – Relevant and Essential

Currently, the only relevant and essential patents that pertain to implementations of PDF/is have been made Royalty Free by the following Intellectual Property statement.

Adobe Systems Incorporated

Patent Clarification Notice Specific to Use of "Portable Document Format: Image-Streamable"

Adobe has a number of patents covering technology that is disclosed in the Portable Document Format (PDF) Specification, version 1.4 and later, as documented in PDF Reference and associated Technical Notes (the "PDF Specification"). Adobe desires to promote the use of PDF as the basis for a file format called "Portable Document Format: Image-Streamable" ("PDF/is") that is currently under development by the Printer Working Group ("PWG"), a program of the IEEE-ISTO.

This Patent Clarification Notice is in addition to the permissions statement set forth in Section 1.4 of the PDF Reference which shall also apply to Adobe's contribution to PDF/is.

Accordingly, Adobe agrees to provide a Royalty Free License to all Essential Claims solely for the purpose of implementing PDF/is. Adobe and the PWG will identify and establish, within the final, published "Candidate Standard" or final "Standard" release of PDF/is, a process whereby implementers of PDF/is can request and obtain the above license.

No license shall be extended to those implementing only draft versions of PDF/is unless that implementation is only used for testing and prototyping purposes.

A "Royalty Free License" shall mean a license that:

i) shall be available to all implementers of PDF/is worldwide, whether or not members of the PWG;
ii) shall extend to all Essential Claims owned or controlled by Adobe and its Affiliates;
iii) shall not be conditioned on payment of royalties, fees or other consideration except as described in (iv) and (v) below;
iv) may be conditioned on a grant of a reciprocal license on identical terms to all Essential Claims owned or controlled by the licensee and its Affiliates; and
v) may include reasonable, customary terms relating to operation or maintenance of the license

relationship including but not limited to the following: choice of law, dispute resolution, and patent notices.

"Essential Claims" shall mean all claims in any patent or patent application, in any jurisdiction in the world, that (A) Adobe and/or its Affiliates own and (B) that would be necessarily infringed by implementation of PDF/is. A claim is necessarily infringed hereunder only when a licensee can prove that it is not possible to avoid infringing it because there is no non-infringing alternative for implementing the required portions of PDF/is. Existence of a non-infringing alternative shall be judged based on the state of the art at the time a licensee implements PDF/is.

The following are expressly excluded from and shall not be deemed to constitute Essential Claims:

1) any claims other than as set forth above even if contained in the same patent as Essential Claims; and
2) claims that would be infringed only by
   a) portions of an implementation that are not required by PDF/is
   b) enabling technologies that may be necessary to make or use any product or portion thereof that complies with PDF/is but are not themselves expressly set forth in PDF/is; or
   c) the implementation of technology developed elsewhere and merely incorporated by reference into PDF/is.

For purposes of the Essential Claims definition, PDF/is shall be deemed to include only architectural and interoperability requirements and shall not include any implementation examples or any other material that merely illustrates the requirements of PDF/is.

An "Affiliate" of a first entity is a second entity that is controlled (greater than 50%) by, in control of, or under common control with the first entity.