

## Media Format Choices for RFCs

draft-crocker-rfc-media-00

### Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as “work in progress”.

The list of current Internet-Drafts can be accessed at <<http://www.ietf.org/ietf/1id-abstracts.txt>>.

The list of Internet-Draft Shadow Directories can be accessed at <<http://www.ietf.org/shadow.html>>.

This Internet-Draft will expire in February 2009.

### Abstract

Documents in the RFC series normally use only plain-text ASCII characters and a fixed-width font. However, there is sometimes a need to supplement the ASCII text with graphics or picture images. The historic solution to this requirement, allowing secondary PDF and Postscript files, is seldom used because it is awkward for authors and publisher. This memo suggests a more convenient scheme for attaching authoritative diagrams, illustrations, or other graphics to RFCs. It further proposes conventions for additional input and display formats, to improve readability. This proposal is based on draft-rfc-image-files-00, by Braden and Klensin, and revises it as little as possible, while expanding the goals of the effort.

## Table of Contents

<b>1. Introduction</b> .....	<b>3</b>
1.1. Goals .....	3
<b>2. A New Scheme for Representation</b> .....	<b>5</b>
<b>3. Construction of the Image File</b> .....	<b>6</b>
<b>4. Requirements for the Base File</b> .....	<b>7</b>
4.1. Overview .....	7
4.2. Figures Section .....	7
4.3. Formatting Changes .....	8
<b>5. Submission and Processing of the Image File</b> .....	<b>9</b>
<b>6. Implementation Issues</b> .....	<b>10</b>
<b>7. RFC Repository File Formats</b> .....	<b>11</b>
<b>8. Internationalization Considerations</b> .....	<b>12</b>
<b>9. Security Considerations</b> .....	<b>13</b>
<b>10. IANA Considerations</b> .....	<b>14</b>
<b>11. Acknowledgments</b> .....	<b>15</b>
<b>Author's Address</b> .....	<b>16</b>
<b>Intellectual Property and Copyright Statements</b> .....	<b>17</b>

## 1. Introduction

Published documents in the RFC series normally use only plain-text ASCII characters and a fixed-width font [RFC2223]. This simple convention has the advantage of a stable encoding for which a wide variety of tools are readily available for viewing, searching, editing, etc.

Inclusion of diagrams, state machines, and other graphics in RFC text has generally relied on the imaginative use of ASCII characters ("ASCII artwork".) However, in a few cases over the years, ASCII artwork has been inadequate for images needed or desired in RFCs. The old solution to this dilemma has been to allow three versions of an RFC: a primary ASCII version and secondary versions that are encoded using PDF and Postscript. The PDF and Postscript versions are "complete", containing a copy of the text as well as the full images [RFC2223]. The textual content and layout of the PDF/PS version is required to match the base version as closely as possible. However, the ASCII text version is considered the official expression of the RFC, and it is always normative for standards track documents. We will refer to this old approach as ".txt+.pdf+.ps" encoding.

The three versions of an RFC using .txt+.pdf+.ps encoding are in separate files in the primary RFC repository (<http://www.rfc-editor.org/rfc/>), with suffixes ".txt", ".pdf", and ".ps". The RFC Editor search engine returns links to all three versions when they are present in the repository.

Unfortunately, the .txt+.pdf+.ps scheme has been awkward for both editor and author, and it is error-prone, so it has seldom been used (roughly 50 out of 5000+ RFCs). The problem is that, in general, only the author has the tools to prepare the PDF and Postscript versions. The RFC Editor edits (only) the primary text version, and then the author must incorporate all the resulting changes into the PDF/PS version while maintaining the "look" of the RFC to the extent possible. There is no practical way for the RFC Editor to verify that this is done correctly, perhaps leading to editorial errors and usually lengthening publication time for these documents.

This memo suggests a much better scheme, for including figures, illustrations, and graphics to an RFC, as well as for maintaining a single copy of base text which can be turned into multiple presentation forms. We hope that the method proposed here will solve the image problem for RFC publication, although the .txt+.pdf+.ps approach would still be possible (and in any case, RFCs using the historic scheme will continue to exist in the RFC repository forever).

This proposal is based on draft-rfc-image-files-00, by Braden and Klensin, and revises it as little as possible. As an expedient, the References section has been omitted from this initial version of the draft.

### 1.1. Goals

The list of goals in the current proposal expands upon the ones of the draft-rfc-image-files-00 proposal:

1. There is a single, master file for document text.
2. Base text is able to be edited, viewed, compared and searched with extremely minimal set of tools, such as a classic text editors, and the like.
3. The master file is subject to formatting constraints, to improve readability when using simple display tools.
4. All formats use strictly open standards.
5. Any mapping from the master file to a presentation format must only depend upon well-tested, reliable tools that are available as open-source.
6. Multiple display forms must be supported, notably scroll-form for screen display and paginated form for printing, as well as classic, basic IETF ASCII paginated format.

7. Figures can be encoded in classic ASCII art and/or in a graphics format.
8. Enhanced display formats can support basic font changes, within IETF-defined criteria, since this can enhance readability.
9. Naming conventions tightly link additional files that are used by the master file.
10. Changes seek to have as much automation as possible for the technical aspects of RFC development and production.
11. Format for the master file should facilitate later revision efforts.

## 2. A New Scheme for Representation

Under our scheme, an RFC may be either a single ASCII file as commonly used today, or a composite of multiple files: an ASCII-only "base file" containing the text of the RFC, and one or more "image files". The ASCII file may optionally conform to xml2rfc format. When present, the image file would be a {{standard image} file that contains only images, captions, and title information. The base file may contain classic "ASCII art" and refer to external image files as alternatives. An RFC which is displayed in any form other than simple ASCII would then be a logical entity whose complete, or preferred, representation could require multiple files, base and image(s).

The base file may be formatted exactly like current ASCII RFCs, with three minor exceptions described below. Alternatively, it may be formatted using xml2rfc. The xml2rfc convention is well-established within the IETF and RFC community. It permits having a single, textual document base, which can easily produce .txt+.pdf+.html formats. In addition, it can contain a text-only version of art, while using external image files, when available and appropriate to the output form.

The intellectual property boilerplate in the base file ("Rights in Contributions BCP 78, RFC 4748 [RFC4748] ) would apply equally to the image file. An image file would contain one or more items that will be known collectively as "figures", whether they are actually diagrams, pictures, tables, artwork, or other non-textual constructions.

This scheme was inspired by the tradition in book publishing, where pictures, figures, or "plates" may be grouped together following the text ("end figures"), or even bound separately from the main body of the text.

In principle, we could allow an image file to be encoded using both PDF and Postscript, since mechanical translation is possible in both directions. However, in the 20 years since the adoption of the .txt+ .pdf +.ps scheme, the PDF format has become a de facto standard for electronic documents, and readers for it are universally available. Furthermore, PDF is being standardized as a format for document archiving, as discussed further in the next section. Therefore, we propose to allow only PDF for image files, simplifying the new approach by not including a Postscript file option.

An ASCII RFC traditionally uses a file name in the form of "rfcN.txt", where N is integer RFC number without leading zeros. The image file that is associated with RFC number N could be named "rfcN.{image name}.{image format extension}". As noted earlier, the repository already contains RFCs with file names "rfcN.ps" and "rfcN.pdf", using the historic .txt+.pdf+.ps scheme.

### 3. Construction of the Image File

Each image would be in a single {image format} file, containing only that image and consistent with the description in [RFC3778] and defined in [ISO32000-1]. The particular {image format} form must be version-stable and must not contain any external references in scripts or otherwise. The RFC Editor authorizes the set of {image formats} that are permitted for use.

There is an issue of whether particular generators of {image format} that claim to satisfy {image format standards} actually do so. Future experience may require published guidelines on PDF-generating software that claims to satisfy {image format}{image format} but does not.

Except as otherwise specified in this document, an image file should contain only a single figure, supporting labels and captions, headers, and footers. It should not contain explanatory text or other materials that could reasonably be expressed in plain-text form in the base file

For xml2rfc output that produces .html or .pdf, images are produced inline and are consecutively numbered.

For .txt output, pages of the image files would be consecutively numbered. The first page number of the image file would follow the last page number of the base RFC, exclusive of the number of the end-of-RFC boilerplate page. The page number of the end-of-RFC boilerplate (in the base RFC file) would be the first page number after those in the image file. Each page of the image file would contain the same headers and footers as the base file, except for one change in the footer, suggested below.

Figures included in the image file would have to be labeled in a fashion that facilitated referencing from the base RFC. They may be numeric and monotonic or it may use textual image names. Simple consecutive integer will usually be the best choice, but in some cases it might be desirable to use a hierarchical scheme like: <section #>.<fig #>. An author who believes that another labeling scheme would increase clarity should check with the RFC Editor.

## 4. Requirements for the Base File

### 4.1. Overview

A base file would be unchanged by the presence of an image file, except for the following.

- For .txt format, the page number of the end-of-RFC boilerplate page would be changed to be logically one page after the last image file page.
- A new unnumbered "Figures" section would be required. This is described below.
- For a composite RFC, a minor modification to the first-page header of the base file and to the footers of both base and image files could tie the additional files together. This is described below.

### 4.2. Figures Section

An RFC that used this scheme (and had any figures) would need to include a Figures section in the ASCII base file. The Figures section should immediately following the Table of Contents, if any, and precede the body of the document. The Figures section should list all figures in tabular form, indicating for each one the figure identification, title, and page number(s).

The style for the Figures section has not yet been fully specified. Here is a suggested example.

---

```

Table of Contents

1. Introduction ..... 1
2. Philosophy ..... 7
  2.1 Elements of the Internetwork System ..... 7
  2.2 Model of Operation ..... 8
  2.3 The Host Environment ..... 8
  (etc)

Figures

Figure 1: Protocol Layering ..... 2
Figure 2: Protocol Relationships ..... 9
Figure 3: TCP Header Format ..... 15, *86
Figure 4: Send Sequence Space ..... 20
Figure 5: Receive Sequence Space ..... 20
Figure 6: TCP Connection State Diagram ..... 23, *87
Figure 7: Basic 3-Way Handshake for Connection
        Synchronization .....31, *88
  (etc)

*Page in Image file

(Page 1 follows)

```

---

An RFC that includes a base file may include ASCII artwork that is suggestive of a figure in the image file, but there is no requirement to do so. When such an approximate figure appears as ASCII artwork in the base file, its figure identification and caption must match those of the corresponding figure in the image file, and the entry in the Figures table should specify the page numbers in both the base and image file. In the example shown above, image file page numbers are marked with an asterisk. Note that very simple ASCII artwork need not appear in the image file.

### 4.3. Formatting Changes

It would be necessary to tie the base and image files together, to make clear they are part of one RFC. Here is an initial suggestion for formatting, which needs further consideration before it is adopted.

The header line "Request for Comments: nnnn" in the base file could be changed to "Request for Comments: nnnn/Base". For consistency, the lefthand footer could become "RFC nnnn/Base". The lefthand footer in the image file could then be: "RFC nnnn/ {Image Name}."

The following sentence could be placed in the "Status of this Memo" section: "This RFC is a composite of this base file and {image format} image files."

## 5. Submission and Processing of the Image File

If a image files are needed, they should be submitted as a `.{image name}>.{image format}` files along with the ASCII text file. The image files should be submitted without headers or footers. The RFC Editor will overlay the image file with the appropriate headers and footers, with correct pagination. The RFC Editor will not normally do any editing of the image file beyond this. If editing the base file reveals problems with figures in the image file, the authors will be asked to create a new image file.

## 6. Implementation Issues

This scheme has a number of implications.

1. The Internet Draft repository must allow submission and retrieval of both base and (when present) image files.
2. Internet Draft file names could be draft-...-vv.txt and (optionally) draft-...-vv.{image name}.{image format}, where "vv" is the normal version number. Updating any file of the composite RFC should increase the version numbers "vv" in both files. We DO NOT want two separate version numbers for one I-D
3. The RFC Editor would need to be able to overlay headers, footers, and page numbers on a given image file. It is claimed that at least Adobe Acrobat Professional includes this capability, and that it also has limited editing capability.
4. The RFC Editor would also need a tool to verify that a given image file satisfies the constraints of {image format}.{image format}.
5. Some RFC Editor scripts and tools would need small extensions.
6. xml2rfc already supports external image files, as an adjunct to, or replacement of, textual art and is used when available, for .html and .pdf formats..

## 7. RFC Repository File Formats

A frequent reaction to the suggestion given in this memo is some confusion over the different file formats that appear in the RFC repository. Here is a brief summary.

If a {image format} image file exists along with a base ASCII RFC, then RFCs in any other format (e.g., complete {image format} files, HTML, or Postscript) remain supplemental, with the reader taking responsibility for assuring that they are equivalent to the base RFC and image file. That arrangement is identical to the relationship between traditional all-ASCII RFCs and supplemental forms: the RFC Editor has never taken responsibility for guaranteeing that the two are identical in content.

The existing .txt.pdf files are not affected by this proposal. The .txt.pdf files are facsimiles of .txt (base files) in PDF, introduced to help Windows users read RFCs online. However, Microsoft has more recently provided an elementary ASCII editor, which probably makes the .txt.pdf files unnecessary in any case.

In summary:

- .txt: ASCII-only file. In old scheme, complete normative file. In new scheme, text part of composite RFC, or stand-alone text file.
- .ps: Old scheme -- a Postscript file that includes figures and whose text is intended to be the same as the normative .txt file.
- .pdf: Old scheme -- a PDF file that includes figures and whose text is intended to be the same as the normative .txt file.
- .{image name}.{image format}: New scheme: image file(s) part of a composite with .txt or xml2rfc file.
- .txt.pdf: Old scheme: Facsimile of corresponding .txt file.

We note that it would be possible to combine the base and image files into a single PDF file, which would have to follow a naming convention to distinguish it from the .pdf case listed above.

## 8. Internationalization Considerations

Our scheme of image files does not, and is not intended to, support character set internationalization for RFCs. It does not allow an author to omit the ASCII text from the base file and instead include the entire RFC text as one (very large) image file.

However, we should note two special cases.

1. RFC 3743 [RFC3743] on internationalized domain names for Chinese, Japanese, and Korean contains a number of examples that may be hard to follow because they can represent those characters only in "U+nnnn" form. An image file could be used that would show the alternative Chinese characters for the examples. This would not diminish either the ability to search the base text or index the document or its readability for those of us for whom reading Chinese characters is difficult, but it should help those who can read them.
2. Suppose that a proposed RFC contained a section derived from Japanese text. The author might put an English translation into that section of the base document, note that the original was really in Japanese, and attach the Japanese as an appendix in an image file. This should raise no difficulties for informative documents. For normative documents, however, the existence of the Japanese original would raise some issues about what was actually authoritative, which is very undesirable.

## 9. Security Considerations

This specifications addresses documentation standards and adding additional flexibility to them. It does not, in general, raise any security issues. However, unless the specifications of this document are carefully followed, the image format recommended, {image format}, may potentially contain external references or scripts that could introduce security problems. The RFC Editor and other publishers should exercise due care to ensure that no such references or scripts appear in the archives.

## 10. IANA Considerations

This document requires no actions by the IANA.

## 11. Acknowledgments

## Author's Address

**Dave Crocker**

Brandenburg InternetWorking

675 Spruce Drive

Sunnyvale, CA 94086

USA

Phone: [+1.408.246.8253](tel:+14082468253)

EMail: [dcrocker@bbiw.net](mailto:dcrocker@bbiw.net)

## Full Copyright Statement

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an “AS IS” basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

## Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org)<sup>1</sup>.

<sup>1</sup> <mailto:ietf-ipr@ietf.org>