

# The RFC Editor Function at ISI

Bob Braden, Sandy Ginoza, Alice Hagens

USC Information Sciences Institute  
Marina del Rey, CA

January 2009

## PREFACE:

The text of this document about the RFC Editor function is based upon the proposal that USC ISI submitted to the Internet Society in 2006. This proposal was to provide RFC Editor services during 2007-2008, with an optional extension to 2009 (this option was approved). Note that the proposal was written during the summer of 2006; many of the proposed tasks have in fact been completed.

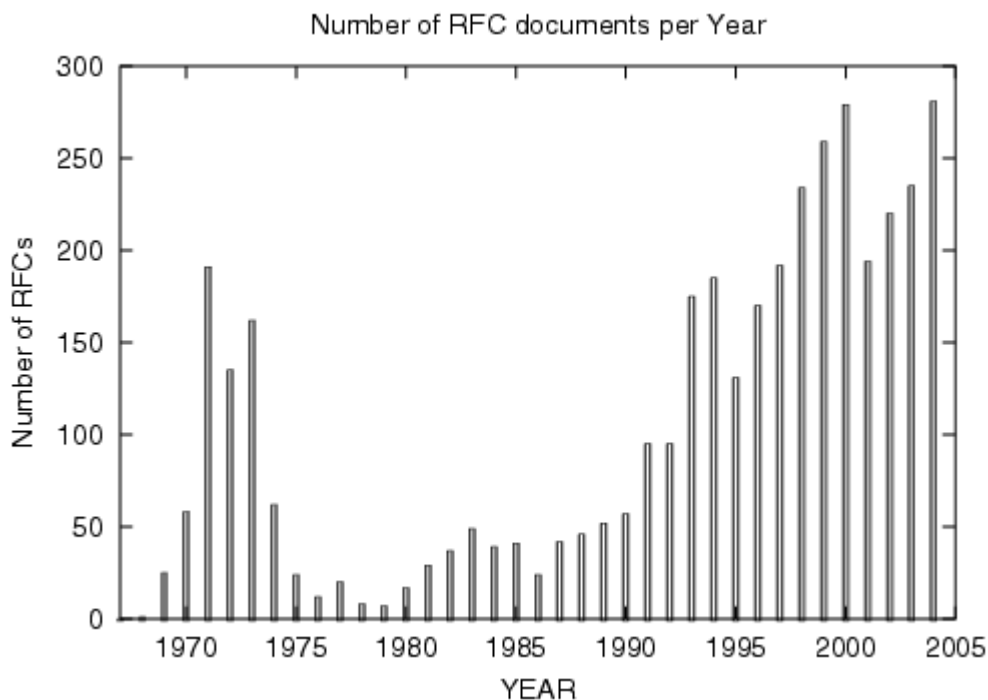
This text is being made public because it contains useful information for the future RFC Editor organization(s), as well as historic information about the recent past of the RFC series. We have omitted budgetary and personnel information as well as various RFP-specific text. It also omits any discussion of the Independent Submission Editor function, which is logically separable from the other RFC Editor functions.

1.	Introduction.....	2
2.	General Considerations.....	3
2.1	Editorial Balance.....	3
2.2	Quality of Published Documents .....	3
2.3	Management and Staffing Issues .....	4
3.	The Current Publication Process.....	5
3.1	Primary Editorial Process .....	5
3.2	Exceptional Cases .....	6
3.3	Submission, Publication, and Archival Formats.....	7
3.4	Supporting Tasks .....	8
4.	Service Levels.....	9
4.1	Input Data.....	9
4.2	RFC Editor Processing Times.....	10
4.3	Base RFC Editor Service Costs .....	12
4.4	Pre-Approval (“Early”) Editing .....	13
5.	RFC Editor Improvements.....	14
5.1	Improvements During 1998 – 2006 .....	14
5.2	RFC Editor Improvements 2007-2009.....	16
5.3	Possible Future Improvements.....	21

# 1. Introduction

The RFC Editor edits, publishes, and catalogs the series of archival documents on computer communication known as Request for Comments (RFCs). For the past 20 years, RFCs have been the official publication channel for Internet standards and other protocol documents produced by the Internet Engineering Task Force (IETF). The RFC Editor and the IETF have complementary and synergistic goals. The function of the IETF is to produce first-class technical specifications, while the RFC Editor's goal to ensure that the published specifications are expressed in clear, correct, and consistent English prose and symbolism and in a consistent and readable format. There should be a strong, collaborative, and mutually supportive relationship between the IETF and the RFC Editor.

The USC Information Sciences Institute, which is primarily a Computer Science research organization, has performed the RFC Editor functions from 1977 to the present. ISI created the current RFC publication process (see Section 3), which published over 5000 documents. The historical record of RFC publication is indicated by the following graph [[http://www.rfc-editor.org/num\\_rfc\\_year.html](http://www.rfc-editor.org/num_rfc_year.html)], which shows the number of RFCs published in each of the last 36 years. Since the average RFC is 30 pages in length, this includes more than 150,000 pages of technical material.



The death of the original RFC Editor, Jon Postel, in October 1998 marked the end of an era for RFC publication, forcing ISI to reorganize the RFC publication services to provide continuity. ISI also began a vigorous program to update the RFC Editor operations and services, as summarized in Section 5.1. The current process is summarized in Section 3; a much more detailed process description will be found in "Procedures Manual for the RFC Editor".

## **2. General Considerations**

### **2.1 Editorial Balance**

In the world at large, the activity of editing often creates a tension between author and editor. ISI attempts to minimize this conflict for RFC publication, while continually striving to produce a uniformly excellent document series. We refer to this fundamental tussle as “editorial balance”, and maintaining this balance must be a continuing concern for the RFC Editor.

The world of technical publishing has generally accepted standards for the typographic rules for “correct” grammar, punctuation, capitalization, sentence length and complexity, parallelism, etc. The RFC Editor at ISI follows these accepted standards, but with particular exceptions. There are a few specific rule variants that have been imposed on RFCs to avoid ambiguity in complex technical prose and to handle mixtures of text and computer languages. There is also a prime directive that must rule over typographic conventions: do not change the intended meaning of the text.

On the other hand, the RFC Editor at ISI respects the long history of individuality in the IETF community. We generally allow variant typography, as long as it is used consistently. Similarly, we allow either British or American English, but if the usage is inconsistent, we will prefer American English. We try to be tolerant of carefully crafted and deliberate alternative styles, although we have our own preference for the "standard" usage. Thus, the RFC Editor at ISI aims to avoid purely “stylistic” changes that, while formally preferable by general standards, do not advance the primary goals of correct English, accuracy, clarity, and consistency. Examples of such “stylistic” changes might be replacing the conjunction "as" with "because", or removing first-person references. In the end, RFC authors and the IESG rule.

ISI has collected editorial guidelines and procedures into a set of documents that are available from <http://www.rfc-editor.org/styleguide.html>.

### **2.2 Quality of Published Documents**

The RFC publication process is a balance between quality, which includes both readability and freedom from errors, and significant customer pressure for short publication delays. It is important to minimize the number of errors in published documents to the extent possible, while maintaining aggressive throughput goals.

The RFC Editor cannot generally be responsible for technical errors, of course. However, correcting purely editorial errors sometimes lead to authors’ discovering and fixing content errors. For example, untangling a tortured sentence sometimes leads to the realization that the original sentence was semantically ambiguous, resulting in the replacement with a sentence that is clear and unambiguous. The RFC Editor also checks for inconsistent use of terminology, both within a single document and among related documents. Correcting such inconsistencies may similarly reveal content errors.

The RFC Editor at ISI instituted a number of management techniques to maintain document quality.

- The Authors' Final Review ("AUTH48") step described below allows authors to reread the edited document and request changes, reducing both editorial and technical errors in published documents.
- The editorial rules and recommendations are documented in some detail, so editing should result in few surprises for authors.
- We created some simple but effective heuristic tools for automating the checking of particular editorial issues. An example is the *matchref* program, which matches citations against references.
- We formalized and automated the errata process using a web portal. The rate of editorial errata reports is a vital measure of error rate in published RFCs, and it is used to alert the editorial staff to specific editorial issues.

### **2.3. Management and Staffing Issues**

Editing is a human-intensive operation that cannot be totally automated. Editing may be difficult and at times intellectually challenging, but it is often tedious, requiring close attention for many hours at a time. Proof reading a dense 120-page technical document and catching a serious typographic or consistency error on page 107 requires great discipline as well as skill. The RFC Editor organization must be able to recruit, train, and motivate a competent and efficient editorial staff.

It takes 2 to 6 months to train even an experienced editor in the editorial rules and conventions specific to RFCs and to the Internet technology. ISI has devoted significant effort to training new people and to upgrading the skills of the current staff. The editorial process can be divided into fairly discrete steps with varying skill levels. For example, we commonly partition the process into three successive phases: (1) copy editing -- marking up the documents using general editorial standards and without specific knowledge about the field, (2) inserting formatting directives, and (3) general editing. The experienced senior editors can of course do all of these, but ISI performs flexible assignment of phases according to the skills of available staff.

The inherent burstiness of the submission of documents is also a major challenge to managing the RFC Editor. One solution used by ISI to handling temporary overloads, bridging inevitable personnel gaps, and training has been to hire junior editors as temporary workers. Since it is not generally possible to train temp editors in the specifics of RFC rules or Internet conventions, ISI has hired "commodity" copy editors who have no specific knowledge about the field. One of ISI's senior editors then goes over each marked-up document and makes only those changes that are appropriate according to the editorial balance guidelines. At this stage, the senior editor also applies the many Internet-specific and RFC-specific rules and conventions.

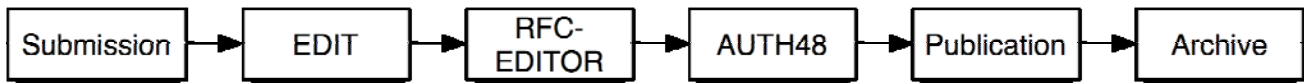
### 3. The Current Publication Process

Some traditional mechanisms for publishing are not appropriate for IETF document publishing, exchanging galley proofs for example. The following describes the process that has evolved for RFCs.

The publication process for an RFC contains a series of stages, which are tracked using transitions among a set of states (<http://www.rfc-editor.org/rfc-editor/rfc-editor-process.gif>).

#### 3.1 Primary Editorial Process

The diagram below represents the major state changes a document moves through during the RFC publication process (see Appendix A for the detailed state diagram). The states were devised to help both the RFC Editor and the community to track the progress of documents in the publication queue.



- Submission

The RFC Editor at ISI currently publishes four document streams: the IETF stream (including both Standards-Track and IETF Informational/Experimental documents), the IAB (Internet Architecture Board) stream, the IRTF (Internet Research Task Force) stream, and the independent submissions stream. Each has its own submission and approval procedure.

- EDIT State

In the editing process, the RFC Editor applies a checklist of some 30 bullets, to maintain consistency and clarity. These include checks for:

- typographic errors (spelling, capitalization, punctuation) or inconsistencies within the document and other documents on the same subject,
- grammar errors and malformed sentences,
- excessively long, tortured, or ambiguous sentences,
- formatting inconsistent with established guidelines,
- inconsistency between citations and references, and
- errors in formal languages (e.g., MIBs, ABNF, and XML).

- RFC-EDITOR State

When editing is complete, the document enters a final quality-control stage, in which the many RFC-specific and IETF-specific rules are checked. At the end of this stage, an RFC number is assigned and inserted in relevant places in the document.

The RFC Editor works closely with the Internet Assigned Numbers Authority (IANA) to register and insert necessary protocol parameters into documents, prior to publication. ISI has long-standing relations with the IANA and has established communication methods to ensure timely processing. The RFC Editor roles include checking for parameter registration requirements that have not been revealed by the IANA Considerations section and inserting the registered values from IANA into the text.

- AUTH48 State - Authors' Final Review

When an RFC is ready for publication, the author(s) are asked to review and approve the final text. Ideally, changes during this stage should be small editorial corrections, not extensive edits or technical changes. However, in practice, problems found at this stage range from trivial editorial changes to significant technical fixes. For the latter, the Area Directors and perhaps the working group become involved and must agree. The senior editing staff is expected to recognize changes that are not solely editorial and require AD approval.

- Publication

When all responsible parties (the responsible parties are specific to stream, see RFC 4844) have agreed, the document is published. This includes putting the publication-format document(s) online, updating index files, notifying IANA of the RFC number for reference purposes (if necessary), and archiving all final source and text files. At this point, the document is announced to the community.

RFCs are published on the RFC Editor website. This site includes hyperlinked access to several indices as well as a convenient search engine. The search engine will return a catalog ("index") entry for one or more RFCs or sub-series documents, matching on title, author, number, or keyword. The RFC Editor also provides access to individual RFCs and to collections of RFCs using SMTP, FTP, and RSync.

### **3.2. Exceptional Cases**

The RFC Editor strives to move documents through the above process as quickly as possible, while maintaining a high level of quality. However, there are a number of possible reasons for significant delay, which also greatly increase the complexity of the editorial task.

- *Normative Reference Hold* - A "normative" reference (i.e., a reference to another standards document) in an RFC must refer to a document that was previously published or must be published concurrently. When a set of related RFCs contain references to each other, all must be held up until they are completely edited and approved, so that they can be published simultaneously. This strict rule resulted from many years of experience with unexpected publication delays that resulted in "dangling" normative references to unpublished documents.
- *Cluster Hold* - Sometimes authors or working groups request that a set of documents be published simultaneously, even when they are not tied together by normative references.

We call such a group of documents (or a group of documents tied by normative references) a “cluster”.

- *IESG Hold* - The IESG may temporarily suspend or withdraw an IETF document from publication to allow further discussion, clarification, or to remand it to a working group.
- *Author Hold* - The RFC Editor may require action by the author for a variety of reasons, technical and/or editorial. For example, the editorial process may have revealed some technical issue or discrepancy, or some change in format may be required that only the author can provide.
- *IANA Hold* - The RFC Editor awaits completion of the actions specified in the IANA Considerations section of the document, or the RFC Editor may have requested action or clarification by the IANA.
- *IESG Expedited Processing Requests* - Such requests from the IESG necessarily delay the documents that are not subject to expedited processing.

These delays add to the complexity of managing the editorial workflow. It might be desirable to provide better communication tools in this area, but these exceptional events are generally outside the control of the RFC Editor.

Normative reference holds in particular have become much more common. To organize the flow of normative-reference clusters, the RFC Editor added two new states: MISSREF and REF (see Appendix A). A document in the RFC Editor queue is in MISSREF state if it is part of a normative reference cluster that includes some documents that are not yet submitted. Once all the documents in the cluster have been submitted, any in MISSREF state move to EDIT state. They each progress through the editing process, perhaps at different rates; the REF state is then used to synchronize and collect the edited members of the cluster to they can enter RFC-EDITOR and AUTH48 states as a group.

### **3.3. Submission, Publication, and Archival Formats**

Currently, ASCII remains the primary format for RFC publication and archiving. This publication text is created by the RFC Editor (and it can be re-created in the future) using the venerable Unix markup tool nroff. Authors of RFCs are free to use any text preparation method. The published ASCII version and the nroff source are archived by the RFC Editor. The RFC Editor may also publish and archive (but not edit) a subsidiary version of a document in PostScript or PDF, but the primary version is ASCII text. In addition, ISI accepts XML source files created for the xml2rfc tool.

The following table summarizes the current formats.

Function	Formats
Submission	.txt or .nroff or .xml
Editing	.nroff or .xml followed by .nroff

Publication	.txt (from .nroff source), .txt.pdf*, .pdf**, .ps**
Archiving	
	Publication formats, .nroff source, and .xml (if submitted)

\* The .txt.pdf is published for each RFC, to help Windows users.

\*\* .pdf and .ps are optional. They are created by the author from the final RFC so that complex diagrams and charts can be included.

**Table: Formats for Submission, Editing, Publication, and Archiving**

### 3.4. *Supporting Tasks*

The RFC Editor performs many tasks that are not directly related to editing and publishing. For example, the RFC Editor:

- Maintains a web page that shows a range of information, including:
  - a listing of the current publication queue,
  - news of significant changes in service or policy,
  - discussion of editorial policies,
  - reports presented to the IETF, and
  - a wide variety of historical data on publication performance.
- Manages editorial review of documents in the independent submission stream.
- Coordinates with the IESG, the IAB, the IETF, and the IAOC/IAD.
- Provides a liaison to attend many IAB teleconferences and meetings as well as IESG teleconferences.
- Reports to the plenary session at each IETF meeting (three times a year) on the status of the RFC process.
- Staffs a “help” desk at IETF meetings with two senior editors.
- Presents a tutorial on RFCs and RFC authoring at IETF meetings.
- Experiments with new procedures and policies that show promise.
- Responds to all email sent to [rfc-editor@rfc-editor.org](mailto:rfc-editor@rfc-editor.org).



## 4. Service Levels

This section discusses RFC Editor service levels, based upon experience at ISI.

Our analysis assumes that the editorial cost of an RFC has a fixed part plus a variable part that is proportional to the number of pages. Here the editorial cost includes all the steps related to specific documents that are performed by RFC Editor personnel.

Note: ISI editors have pointed out that editing time actually has a quadratic dependence upon the number of pages, resulting from cross-checking section numbers, figures, citations, etc. We have no measure of this, and will consider only the linear dependence. The quadratic effect is probably unimportant for documents shorter than 100 pages, and longer documents are relatively rare.

### 4.1 Input Data

- **Document Processing Time**

The cost of proof-reading/copy-editing steps in particular is directly proportional to the page count. The RFP Figure “One Copy-Editor Documents and Pages: October 2005 – June 2006” shows that ISI gave the copy editor approximately 300 pages of work each week. By his testimony [private communication], this results in a roughly constant workload of around 30 hours each week. The results shown in the RFP also support a simple linear dependence of copy edit time on page count.

This was a highly experienced copy editor, so this we may realistically assume that copy editing (marking up a hard-copy document) or an equivalent proof-reading process can at best be roughly 10 pages per hour. An independent measurement from another copy editor, essentially matches this result. However, copy editing is only the first stage in the editorial process. Full editing has been measured at 4 pages per hour (7.5 hrs/30 page doc), for the majority of documents that can be efficiently edited. However, there are some outliers whose complex issues or poor English quality require significant additional editorial time. We estimate that 1/3 to 1/2 of the overall average editorial cost can be due to these outliers. Process improvements such as early editing may reduce the number of outliers, but there will always be a significant set of them.

- **Average Page Count**

ISI’s measurements show that the average RFC length is close to 30 pages, although the RFP Figure referenced above shows that the standard deviation of document sizes is large, ranging from 1 to 300 pages (note the single 300 page RFC that took the copy editor the entire week of 4/23/2006). In the following, we will consider document aggregates assumed to have the historical distribution of page sizes, and we will use the term “document” to mean a canonical document of 30 pages. Thus, a 300 page RFC represents 10 documents in our calculations.

The RFP specifies (converting to page-equivalent canonical documents per month)

- In 2007, 440 docs with average size  $P=30 \Rightarrow 36.7$  docs/mo.
- In 2008, 480 docs with average size  $P=33 \Rightarrow 43.2$  doc/mo, or 518 canonical 30-page documents.

Note: We assume the RFP figures on page sizes, but we do not have evidence to support the increasing page length over time.

## **4.2. RFC Editor Processing Times**

The processing time that is under RFC Editor control is the sum of time in the two states EDIT and RFC-EDITOR. Call it EdPT (Editor processing time)

(The validity of such a measure assumes that the editors are scrupulous about tracking state changes from day to day. ISI has regarded such perfection in tracking to be a waste of valuable editorial resources, and so the historic data is only approximately correct.)

Requirement:

By July 1, 2007: 60% of published RFCs will have EdPT < 20 calendar days

By October 1, 2007: 75% of published RFCs will have EdPT < 20 calendar days

By January 1, 2008: 90% of published RFCs will have EdPT < 20 calendar days

We will now outline our approach to estimating the resources needed to meet these goals. It is important to understand this approach, as it contains some (necessary) assumptions and approximations. We start by computing the average processing load over the year, ignoring effects of bursty submission. Then we compute an approximate inflation of this load to account for the burstiness.

### **Average Processing Load**

The most basic requirement is editorial capacity to handle the average document load. Taking into account the projected increase in page numbers, we estimate the annual submission rates for publication to be 440 and 518 canonical 30 page documents in 2007 and 2008, respectively.

Based upon ISI's historical experience 2001 – 2006, we projected editorial hours per document. They formed a roughly decreasing time series from 18 hrs/doc (2001) to 15 hrs/doc (2005). Note that this period encompassed a doubling of the editorial staff size and many changes of editorial staff composition. The consistency of the results implies some robustness in this measure.

The proposed budget level is based upon 15 hrs per document per Editorial FTE. The difference between this (measured) rate and the 7.5 estimated earlier is due to outliers that require exceptional effort and to tasks that are indirectly related to document processing but are not specifically editing, such as updating the database.

## Percentage Limits

As noted above, the RFP requirement limits not just the mean, but rather a measure of the distribution. This is a much harder problem than the mean. Because of the large burstiness of the submission rate (see the IETF Community Submissions chart in the RFP), the RFC Editor processing must be modeled as a classical single queue, single server system. Developing a believable distribution of arrival times is difficult, so we used the actual distribution.

We extracted ISI historic data on documents submitted to the EDIT state, between January 3 2005 and March 13, 2006. The total was 458 documents containing 13068 pages, for an average of 30.0 pages per (calendar) day, and 38.5 pages per document.

We wrote an awk script to calculate the complete queue history in documents and pages, using this historic data on arrivals and assuming a constant service rate.

Although the submission times were real, the calculated publication times bore no relation to reality, because they ignored the many holds caused by exceptional cases. However, total processing times are discussed below. The following table summarizes the results.

Capacity (ppd)	# Docs Published	# docs left in Q	Avg days in Q	% < 20 days	# Pages Published	Avg Pages Per day
10	143	315	156.7	4%	4352	10.0
20	312	146	88.3	10%	8715	20.0
27	420	38	24.5	36%	11613	26.6
28.45	421	37	17.5	60%	11808	27.1
29	422	36	15.4	67%	11848	27.2
29.85	424	34	12.9	75%	11906	27.3
31.70	426	32	9.6	90%	12010	27.5
35	435	23	7.0	98%	12042	27.6
50	446	12	2.7	100%	12432	28.5

This shows that the service capacity (in pages per calendar day) required for 90% compliance is approximately 10% higher than the capacity for 60% compliance, and 6% larger than the mean publication rate required. In the staffing and budget plan in Section 9, the staffing in 2008 is inflated by approximately 10% to meet the 90% requirement.

***A key assumption is that the distribution of input rates and pages sizes in 2007-2008 will roughly correspond to those in 2005-2006 (the means can change, without changing the conclusion).***

An approximation in the simulation was to use calendar days; the existence of weekends actually causes some additional burstiness, but we assume that this will not affect the conclusions.

### 4.3 Base RFC Editor Service Costs

The earlier discussion leads to an estimate of the editorial staffing for 2007 and 2008, shown in the following table. The second column is the assumed productivity. This is given in hours per document per editorial FTE (eFTE). This time includes *all* of the many tasks involved in editing and publishing a single document, and is based upon historic data at ISI. Although we expect some additional efficiency will arise in 2008 from the improvements in 2007, we anticipate counter-balancing additional costs for the additional services required by the RFP (see Section 6.1). We therefore assume 15 hours per document per editorial FTE in both years.

The third column converts this to documents per month per eFTE, assuming 140 working hours per month. The fourth column repeats throughput required by the RFP in 30-page-equivalent documents per month. A division yields the last column, the number of editorial FTEs required each year.

Calendar Year	Hr/doc (see Section 6.2)	Docs/mo/eFTE	30 pg Docs/mo (see RFP)	Editorial FTEs (eFTE's)
2007	15	9.3	36.7	3.9
2008	15	9.3	43.2	4.6

Notes:

- The difference shown between FTEs and editorial FTEs is the staffing of the functions called for by the RFP that are not directly related to editing and publishing specific RFCs: RFC Editor management and administration, development of improved procedures and services, liaison and coordination with third parties, creation and maintenance of user documentation, response to email questions, etc.
- The total 2008 Edit FTE includes the approximate 10% inflation implied by the queue simulation in Section 6.2.2, to provide 20-day turnaround for 90% of the documents.

These conclusions are based upon the following assumptions:

- The specified total document loads in each year.
- The specified mean page sizes in each year.
- That the distribution of input rates and pages sizes in 2007-2008 will roughly correspond to those in 2005-2006 (the means can change, without changing the conclusion).
- That the quadratic term in editorial load as a function of document size can be neglected.
- That the historical experience of ISI is a reliable predictor of the future, taking account of expected improvements in efficiency.
- That the queue simulation (Section 15.1) based upon 2005-2006 submission timing provides a good indication of the standard deviation of processing times, scalable to other document loads. In particular, we assume that ignoring weekends makes little difference.
- That the community will not add any additional tasks or significantly change the nature of the specified tasks.
- That the early (“pre-approval”) editing service, if contracted for, will be in conjunction with the base effort.

Finally, this discussion of service levels does not include the effect of differing priorities among different RFC streams or the effect of expedited publishing requests

#### **4.4 Pre-Approval (“Early”) Editing**

The current post-approval editing system, in which documents are edited only after they are approved, provides the most efficient use of RFC Editor resources, because document queuing smoothes out the substantial burstiness and keeps all editors optimally busy. On the other hand, ISI recognizes the inherent advantages of doing editing earlier, while the documents are still in the working group process. ISI therefore participated in an experiment with early editing, in which an editor was assigned to read and suggest changes in documents that were still in preparation in the working group.

It was clear that the working groups preferred the early resolution of issues in the document. However, the early editing experiment did not produce conclusive evidence on overall editorial efficiency with this system. We believe that early editing + post-approval editing will require somewhat more resources than pure post-approval editing, but we have insufficient data to understand the cost or the balance of resources it implies.

We considered two scenarios for early editing (others are possible).

- Copy editing

This scenario would provide a professional copy editing service, producing a marked-up hard copy of an Internet Draft, to clean up the English. This copy editor might not have more than limited knowledge of Internet-specific and RFC-specific editorial conventions, and he/she might suggest some “stylistic” changes that conform to common publication practice but are beyond those called for by “editorial balance”. Based upon ISI’s historical experience, we estimate this scenario at 10 pages per editor hour.

- Full editing

This scenario would provide an early version of the comprehensive editing currently performed on approved documents, including checks for consistency and clarity and with knowledge of Internet conventions and RFC-specific rules. It would require the use of xml2rfc to prepare the documents. The results would be provided in soft copy as a modified xml2rfc source file. Based upon ISI’s experience, we estimate this scenario at 4 pages per editor hour.

We believe that the full editing option is much preferable for the IETF. However, without data on the burstiness of demand for early editing, it is not possible to predict how much manpower it will take to keep the processing time for early editing under 10 days, for example.

## 5. RFC Editor Improvements

### 5.1 Improvements During 1998 – 2006

While Jon Postel was RFC Editor, his procedures and policies included many historical aspects that had accumulated since the series began in 1969. Since 1998, the RFC Editor at ISI has also worked to replace or remove obsolete mechanisms and conventions, updating the RFC Editor function while retaining the essential features. These changes were made carefully and incrementally, but the net result was a large shift towards efficiency and transparency of the operation. Many of these changes were in response to suggestions and complaints from members of the Internet technical community. This section samples these post-1998 changes in the RFC Editor function.

- Improved Transparency
  - ISI completely revamped the RFC Editor web site, to include convenient access to search engines, alternate views of the RFC archive and index, instructions to users, policies, and news.
  - ISI formalized the state diagram [<ftp://ftp.rfc-editor.org/in-notes/rfc-editor/rfc-editor-process.gif>] for the publication process and updated it to more accurately track document progress.
  - ISI created a user-friendly search engine for RFCs and Internet-Drafts in the ISI repository. This repository is primary for RFCs and mirrors the IETF site for Internet Drafts. The RFC search engine shows STDs and BCPs as well as RFCs, and it shows obsoleted documents in a distinct font.
  - ISI installed automated email to authors (cc'ing working group chairs and area directors), so that whenever the publication state changes, state-dependent messages are sent. This provides authors with relevant state change information and automates routine email messages.
  - ISI implemented an authors' 48 hours notice message that gets cc'ed to working group chairs and area directors. Reminders are sent to authors on a weekly basis, and a message goes to ADs (cc'ing authors and WG chairs) when there has been no response.
  - ISI created a script to produce a daily summary of all documents in the RFC Editor queue. This report is used by ISI staff to track documents, and it is supplied weekly to the IANA and to the IETF and IAB chairs.
    - Each document that requires IANA processing is shown with a “\*A” flag.
    - Each document that has unpublished normative references is shown with a “\*R” flag.
- Improved Services
  - ISI promoted the Abstract into a first-class part of every document. The Abstract can be displayed in the search engine on the RFC Editor web site, for example.

- ISI created an archive of errata for published RFCs. This list is on the web site and is linked to the search engine, so that search results include hyperlinks to any errata items.
  - ISI instituted the use of `htmlwdiff` and sends its output to authors to highlight editorial changes.
  - ISI created a secondary archive of RFC documents for the convenience of Windows users: PDF facsimiles of each ASCII RFC. They exist in the archive with file names of the form: `rfcnxxx.txt.pdf`.
  - ISI experimented with and subsequently adopted the use of `xml2rfc` in the editing process. We now accept the XML as a submission format (along with the corresponding `.txt` file) and use it to improve editing efficiency when possible.
  - ISI collaborated with the IESG to conduct an experiment in “early” editing of documents while they are still in the working group process.
- **Improved Coordination**
    - ISI worked with IANA to clarify and improve synchronization of IANA protocol parameter assignment with editing. In particular, we modified our procedures to allow parallel processing with IANA assignment.
    - ISI reorganized publication states to clarify the impact of normative reference holds. This involved two additional states, `MISSREF` and `REF`. The `MISSREF` state contains documents that contain at least one normative reference to a document that is not yet submitted to the RFC Editor. The `REF` state contains documents that contain at least one normative reference whose editing is not yet complete.
    - ISI established an Editorial Board to advise the RFC Editor on independent submissions as well as general editorial policies.
  - **Improved Efficiency**
    - ISI worked with the `xml2rfc` development community to make `xml2rfc` an effective tool for RFC publication. The `xml2rfc` source language was primarily designed for ease of document preparation, so it did not provide the fine control over formatting that is required for final markup for publication. However, ISI was able to make effective use of `xml2rfc` source to do the great majority of editing, but at the final stage prior to publication we convert the source from `.xml` to `.nroff` source to perform the fine-tuning of format.

This proposal includes further efforts to maximize the power of `xml2rfc` for RFC publication, and we anticipate significant productivity gains as a result.

- ISI created an online database of RFC reference entries, to speed up processing. This is also available to authors via the RFC Editor web site.

- Editorial Procedures and Policies
  - ISI improved the consistency and accuracy of the editorial process. For example, ISI:
    - Set editorial guidelines for abstracts, titles, Tables of Contents, and abbreviations.
    - Instituted formal language checking for MIBs and XML schemas as well as ABNF.
    - Instituted checking that references to IETF documents are the latest versions.
    - Created tools for routine checking of references and formatting.
  - In collaboration with the IESG, ISI designed and implemented the division of references into "Normative" and "Informative".
- System Changes
  - ISI replaced the historical "flat-file" RFC index database with a MySQL database.
  - The HTML file that contains the current publication queue [<http://www.rfc-editor.org/queue.html>] is now generated automatically from the index database. This eliminates typographic errors that used to creep into the queue file.

## **5.2 RFC Editor Improvements 2007 – 2009**

ISI proposed the following improvements in support of quality, efficiency, and transparency. Several of these improvements (**called "innovations" in the RFP**) involve new tools and the use of technology for enhanced communication, as well as more consistent training of the editorial staff.

We classify these improvements into 3 categories: Quality, Efficiency, and Transparency.

### **1. Innovations to Improve Quality**

- A - Responsible Editor
- B - New RFC-EDITOR (Quality Control)
- C - Editorial Meetings
- D - Statistics

### **2. Innovations to Improve Efficiency**

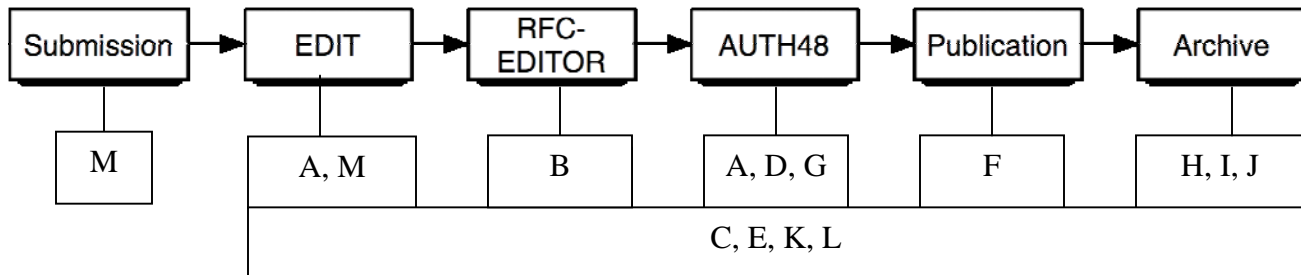
- E - Increased Use of xml2rfc
- F - Automated Publishing

### **3. Innovations to Improve Coordination and Transparency**

- G - New AUTH48 (Author Approval)
- H - Errata Portal
- I - RFC Online
- J - XX99 & XX00 Series Documents
- K - Queue Display
- L - Web Services Interface
- M - Website Navigation and Content



The diagram below shows the current process flow for all IETF Internet Drafts submitted for publication. The letters indicate where the innovations (A-M) affect the publication process.



ISI intends to implement the listed features and commits to examining the process on a continual basis to adapt to the changing needs of the community. Section 5.3 highlights several possible future improvements.

- **Innovations to Improve Quality**

#### **A - Responsible Editor**

After a document has been copyedited on the hard copy, the editor who inserts edits into the soft copy will become the single “responsible editor” for the document. This editor communicates with authors and is familiar with the issues particular to this document. This editor is responsible for the document’s progression through the queue until final publication. They will read and filter through the copy editor’s marks, insert edits as appropriate, and dialog with the authors (if necessary) to resolve any queries and/or problems. Once the editor has completed the EDIT tasks (3.1), the document will go to RFC-EDITOR check (see B, below).

This responsible editor will also work with the authors during the AUTH48 stage. Using a single, continuous editor for each document allows a document to remain with an editor who is already familiar with the work and author preferences. The author is able to provide direct feedback to the individual working on the document. This continuity in workflow should reduce the documents total queue time, as AUTH48 changes should be decreased.

#### **B - RFC-EDITOR (Quality Control)**

See Section 3.1 for a description of the current RFC-EDITOR state. We plan to adopt a new peer-to-peer approach for this state that will distribute the quality-control step across the senior editorial staff.

- Once the responsible editor has completed their editorial processing, the document will be passed to another member of the staff for RFC-EDITOR check before AUTH48.
- This review involves examining the diff file and checking many areas, including IESG notes, IANA actions, author requests, reference problems, checklist items, and editorial balance.

- An RFC number will be assigned and cross references updated.
- Note that this quality-control process will often deal concurrently with all the RFCs in a “cluster”, to be published together .
- At completion, the AUTH48 notice(s) for the RFC (or set of RFCs) will be sent out.

We believe this approach increases uniformity and quality in the document series. This requires all the senior editors to be cross-trained to perform the quality-control process well. Comments and editorial issues will be collected for review and discussion at weekly editorial meetings (see C).

## **C - Editorial Meetings**

Formalizing this aspect of continual training at the RFC Editor, we hold weekly meetings to review editorial issues. These issues arise from 4 sources: (1) questions during EDIT state, (2) feedback from the altered RFC-EDITOR state, (3) AUTH48 changes from authors, and (4) reported errata. The goals of these meetings will be to consolidate RFC Editor style and editorial practices. After each meeting, editorial policy points are recorded.

## **D – Statistics**

ISI proposes to keep statistics on the number of rejected edits during the AUTH48 process. We will include a notice in the AUTH48 notifications that requests authors indicate rejected changes. We define rejection of an edit to be reversion of the text to the original (as it was submitted in the ID). We will report on these findings on a monthly basis..

In our experience, it is important to identify and review rejected changes because these are edits for which we have changed the intended meaning. Each of the rejected changes will be discussed at weekly editorial staff meetings (see C). In addition, documents requiring a high number of changes during AUTH48 also will be flagged for discussion at weekly editorial staff meetings. The number of rejections will serve as an indication of our EDIT performance and provide continual training to the editorial staff in making appropriate edits.

We don't feel that a rejection of RFC Editor working for other new wording should be considered a “rejected edit” since the result is new, and presumably improved, text in the document.

- Innovations to Improve Efficiency

## **E - Increased Use of xml2rfc**

Currently, .xml source submitted by authors is edited, then used to generate an nroff source file, which is further edited and updated during AUTH48. We will increase use of the .xml source for handling text changes during AUTH48 by sending it to the authors for insertion of their changes, with the ultimate goal of accepting and archiving .xml as an official source file. In addition, we will continue to work towards using the edited .xml source file to create the final RFC by continuing to provide feedback to the xml2rfc development community. This goal is desirable because it allows authors to have an XML source file that is consistent with the published text, as well as a starting point for writing a revision of an RFC (i.e., an rfcXXXXbis document).

## **F - Automated Publishing**

New scripts will automate additional routine checks for typographic consistency and for clerical tasks such as sending email announcements.

- Innovations to Improve Coordination and Transparency

## **G - AUTH48 Portal**

For increased transparency and better author and editor communication, we will create an AUTH48 web portal that will serve as a checklist for the authors and RFC Editor. The portal will list the authors and will be updated to show which authors have given their approval for publication. This will allow the community to track the progression of the author and editor interaction during the final stages of publication.

This page will also list normative reference issues and their status, respectively. Often, a document set moves to the AUTH48 stage together, but author sign-off does not occur concurrently. The documents in the set are held until all of the documents are ready to be published simultaneously. This portal will provide information about which referenced documents have completed the AUTH48 process and which have not. The listed references will be links to their respective AUTH48 pages.

## **H – Errata Portal**

ISI will institute a new process for handling errata. It will allow immediate errata posting and a streamlined verification procedure using a web portal. This will address the need for giving precedence to verifying significant errata, because the verifiers (RFC Editor, authors, and IESG) will have the option to verify an erratum as soon as it is reported.

Any user will be able submit an erratum via a web form on the RFC Editor webpage.

- The erratum will automatically be posted online and marked "not verified".
- An automatic email notification will be sent to the authors and the IESG.
- The IESG will receive a password for verification purposes.
- Each error will be logged separately; some will be valid, verified, and posted, while others will be discarded.
- Upon verification, an erratum will automatically be added to the main errata page and linked from RFC search engine results.

We propose a verification structure such as the following:

Editorial error -> verification by RFC Editor, ADs, or authors.

Technical error -> verification by ADs (Standards Track) or by authors.

## **I - RFC Online Task**

RFCs that are not online (approximately 200 documents earlier than RFC 800) are currently at ISI in hardcopy. They will be scanned and posted as PDFs. Their information will be added to the database, and they will be made available through the search page. ISI already has the software to do this.

## **J - XX99 & XX00 Series RFCs**

These documents have become obsolete and their publication will be discontinued. The Official Standards Protocol document is a living document (<http://www.rfc-editor.org/rfcxx00.html>) that is updated daily. This will be implemented for the XX99 summaries as well. These files will be made available online and updated dynamically. They will periodically be captured and stored as archive files, but they will no longer be published as RFCs. The online pages will be altered to provide the contextual information necessary to understand the listings.

## **K - Queue Display**

The current queue is displayed on [queue.html](#) in date order and divided by category. The queue page will be reorganized to offer more transparency of the RFC Editor process. There will be a "Detail" field that indicates who the token-holder is and what action is required.

RFCs processed by the RFC Editor frequently come in clusters that need to be published simultaneously. This clustering may be formal, resulting from (chains of) normative references, or it may be informal (see "Cluster Hold" in Section 3.2). Clusters have ranged from 2 to 10 documents in size, and keeping track of them is a significant manual task for the current RFC Editor. In the process of designing and installing improvements in the RFC Editor database (document tracking) system to meet the other objectives of this proposal, we will incorporate a new mechanism to track and display clusters.

## **L - Web Services Interface**

The RFC Editor will use web services technology to provide real-time access to archive meta-data, replacing the always fragile technique of "screen-scraping". Specifically, web services interfaces will be developed and published allowing external services to query about documents pending publication, such as the current state or token holder, and documents in the archive, such as errata, standard level, or download URL. Furthermore, web services interfaces will allow straightforward integration of RFC archive meta-data with IETF and external search tools. We will also investigate the feasibility and desirability of using web services for document action or errata submission. Our expectation is that web services will allow integration of RFC processing status into the IETF Secretariat's ID Tracker tool suite. We will work with the community, in particular the IESG, to ensure useful interfaces are created and then the IANA and Secretariat support staff to ensure smooth integration.

## **M - Website Navigation and Content**

The web pages will be reorganized for easier navigation and more effective communication of the information for authors. A page titled "Information for Authors" would gather links for the following into one place for easier access:

- 1) RFC conventions and policies including items such as formatting rules and style guides.
- 2) Authoring tools (an updated version of <http://www.rfc-editor.org/formatting.html>).
- 3) Checking tools for authors, including the ABNF checker and the script for verifies that references and citations match.
- 4) The reference entries for all RFCs formatted in the style of the RFC Editor (currently at <ftp://ftp.rfc-editor.org/in-notes/rfc-ref.txt> ).

These pages will serve as a common reference for the IETF community and RFC Editor staff. By posting and updating this information (in response to community feedback and editorial meetings), we will maintain a uniform document series and increase transparency. This will also allow a quicker publication process, because it will decrease questions from authors and allow the RFC Editor to send pointers, rather than generate individual replies.

### **5.3 Possible Future Improvements**

ISI's long experience with running the RFC Editor function leads to many useful ideas that we do not have the resources to implement, or which require some experimentation before they can be adopted. This section lists some of those ideas for desirable improvements, as a record for future planning and to demonstrate our desire to innovate in a responsible manner. This is an RFC Editor wish list.

We plan to actively experiment with publication practices in an attempt to find a more efficient means of producing quality documents in a timely manner. A few areas where we see potential for experimentation are described below.

#### **Pre-Approval Editing Option**

Based on our experience with the early copyedit experiment in 2005, we can offer the option of pre-edit before WG last call. This enables authors and working groups to review the suggested RFC Editor changes, and submit a revised document to the Area Directors for approval. The WGs notify us when the document is ready for review, and it is assigned to one editor who will edit the document and contact the authors with any questions.

The editor will send the edited source file to the authors. Then, the authors have the opportunity to revise the text and submit a new version for approval. After the document has been approved for publication and the RFC Editor receives a document action from the IESG secretariat, the document will be returned to the originally assigned editor to review changes made since the previously reviewed version.

We do not currently have enough experience with pre-approval editing to determine the benefits. However, we believe that the pre-edit phase can have a significant impact on the RFC Editor process. For this reason, we support the idea of the early edit process, should the community decide to adopt it.

- **Image Files**

ISI would like to experiment with the publication process by giving authors an option to submit a supplemental file that contains diagrams.

The Internet community has long debated how to include complex diagrams and graphs in RFCs. Out of all the ideas discussed, one stands out as a feasible and useful approach: John Klensin's idea of an "ASCII with pictures" model. In brief, a published and archived RFC could consist of two files: an ASCII text file containing all the text of the document, and a "supplemental" graphic file (probably using PDF) containing figures, diagrams, and other graphic material that cannot easily be represented in ASCII. The ASCII text file would be subjected to the entire editorial process as today and as projected

in this proposal. The supplemental graphic file would be passed through the publication process as a read-only object, not subject to editing by the RFC Editor. We would like to experiment with this, and if it succeeds, institute it as a regular procedure for RFC publication. It is a compromise that appears both feasible and highly useful, and ISI proposes to experiment with it and if possible institute it.

To quote from Klensin's message:

“The idea is to have a supplemental file rather than abandoning ASCII for some or all documents in order to permit pictures. This involves some rethinking of how we maintain our archives, but closely resembles the old book-publishing convention of putting all of the photographs into a single insert that may or may not be bound in. One might then have RFCnnnn.txt and RFCnnnn-supplement.pdf, with the former referencing embedded figures as, e.g., Figure 1, Figure 2, etc and figures in the supplement as Figure S.1, Figure S.2, etc. This approach eliminates the searching and indexing issue that is often cited as justification for retaining an ASCII-only format, since the supplement consists only of figures. Clearly, there are issues that this approach does not resolve such as embedding non-ASCII characters into the text. However, it would clearly resolve the issues associated with the perceived requirement to include line drawings, and even more complex images or photographs, in RFCs.”

As an example of a downside, the RFC Editor will need to verify consistent notation between the supplemental file and the base ASCII text. We need to determine the cost of this in practice.

Note: this approach is not intended to allow authors to slip arbitrary text into the supplemental file to avoid editing.

This approach has been documented in Internet Draft **draft-rfc-image-files-02.txt**.

- Designated Author

This innovation would simplify the relationship to multiple authors on a single RFC. It would ask that one or two out of the list of multiple authors on one RFC be designated as the responsible points of contact for the RFC Editor. The designated author(s) would collect and submit all AUTH48 changes, verify correct incorporation, and give final publication approval. This would allow shorter publication times.

Implementation of this procedural change would require the cooperation of the IESG and working group chairs, as we would need them to enforce this policy. The IESG secretary would include a statement in the Document Action that names the responsible author(s).

- Remote Access Library Facility

To begin the AUTH48 author's review process, the RFC Editor sends the author(s) the edited text and an htmlwdiff file showing the changes. This is accomplished by emailing a URL to the document in a public directory (in-notes/authors/). The author(s) then ask for changes via email in a somewhat clumsy OLD/NEW format. This works adequately in simple cases, but the AUTH48 process sometimes includes several rounds of changes, involving multiple authors and an AD. This suggests (but does not conclusively prove) an advantage from using a remote-access library facility, an example of which is the well-known CVS. This facility would allow authors to directly update the RFC Editor's document, and it would keep track of versions and make it easy to take diffs. Some experimentation on this approach

would be essential to ensure this approach is effective and efficient and has no unintended consequences.

Security is an issue here. Authors must be authorized to update a particular document. A password could be included in the AUTH48 message, or there could be an RFC Editor user registration procedure that would include a login password.

- User Registration

A general-purpose user registration capability could be installed for RFC Editor activities. Since its email address is so widely known, the RFC Editor is prey to spam bombardment. If users (e.g., anyone in the community who needs to communicate with the RFC Editor) were registered, unregistered user email could be handled at lower priority. Registration could also be used to control access to the remote-access library facility, as discussed immediately above.

- Internationalization

We would like to explore the issues of extending the character set beyond US-ASCII, i.e., UTF-8. A major issue is whether there is a set of preparation, display, and searching tools for both the RFC Editor and the RFC consumers. The RFC document series is an international publication that should adopt the international character set.

Appendix A: RFC Editor State Diagram

# RFC Editor

## Document Processing

