2020 Self-Review: RFC Production Center

The overarching goal of the RFC Production Center (RPC) is to produce high-quality RFCs in a timely manner. In 2020, the RPC continued working in the v3 format and, with the developers and others, we reported bugs and noted recommended practices, while also decreasing the overall queue size and related document processing times. While the main focus was on the aforementioned, the RPC also undertook other tasks, as described in the "Other Updates" section.

This self-review will discuss the queue throughput rates for 2020 and the impact of v3-related work, identify other areas in which the RPC has made progress, examine challenges and areas for improvement, and provide some areas of exploration for the coming year.
This self-review will discuss the cause for the increased queue size over time and the agreement to suspend the SLA through most of 2020, as well as the non-RFC-specific work associated with documenting best practices for v3 XML and other tasks. See https://www.rfc-editor.org/report-summary/ for details about the SLA.

Editing and Publishing RFCs

2020 was both a challenging and rewarding year for the RPC. While the switchover to v3 XML took place in September 2019, the RPC continued to deal with the effects of the transition throughout 2020. Simultaneously, the queue was growing, as a number of documents were released from MISSREF to the EDIT queue on top of document approvals in Q3 2019; Figure 4 indicates that 89 I-Ds moved into the EDIT queue in Q3 2019 alone. In addition, a high number of Internet-Drafts (I-Ds) were approved for publication in Q1 of 2020, as 65 I-Ds moved into the EDIT queue. Figure 1 highlights the spikes in the number of documents moving into the EDIT queue during these quarters. Figure 2 highlights the queue size and pinpoints the transition to v3 and when C238 (50-document cluster) moved into RFC-EDITOR state. This combination of events resulted in a significant increase in processing times throughout most of 2020. As the relevant leadership bodies and community understood the undertaking associated with the v3 transition, the SLA was deferred for most of 2020.
Note: Please see <https://www.rfc-editor.org/about/queue/> for background and state definitions.

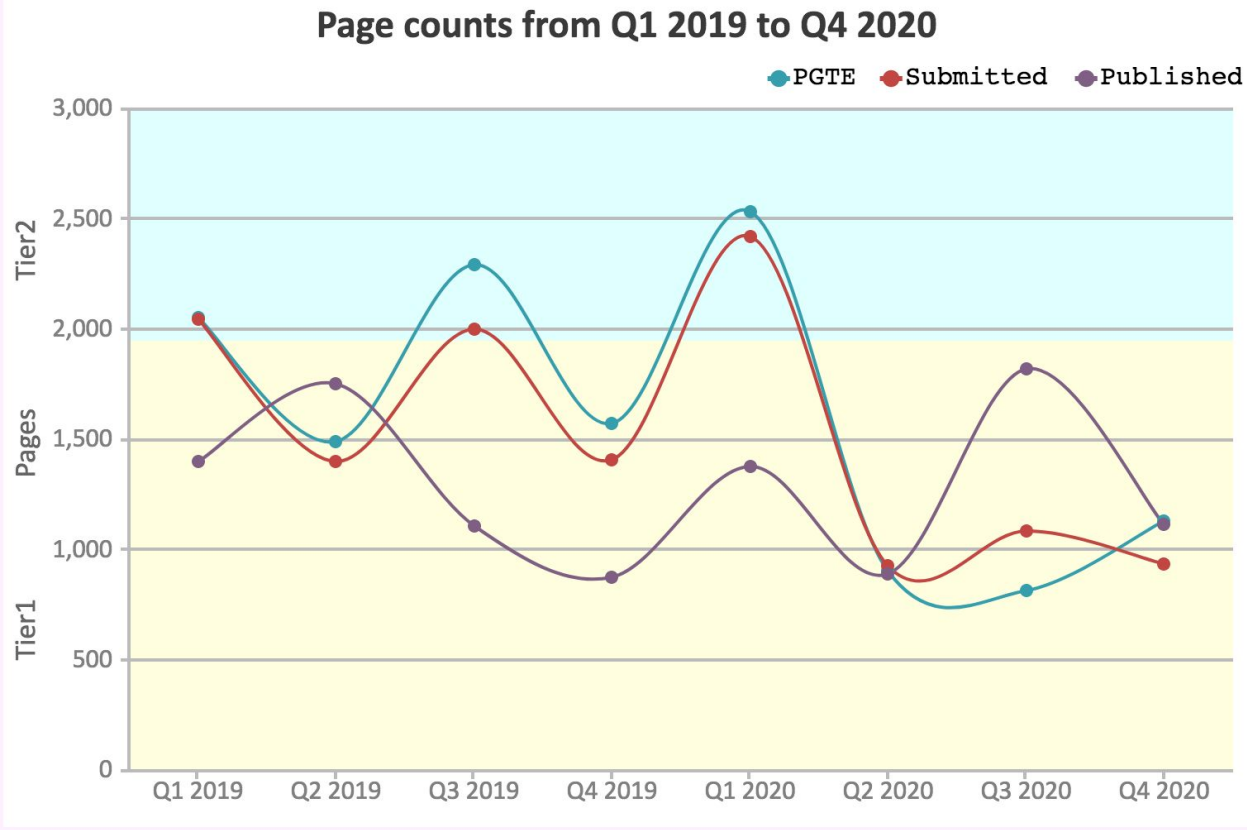

Figure 1

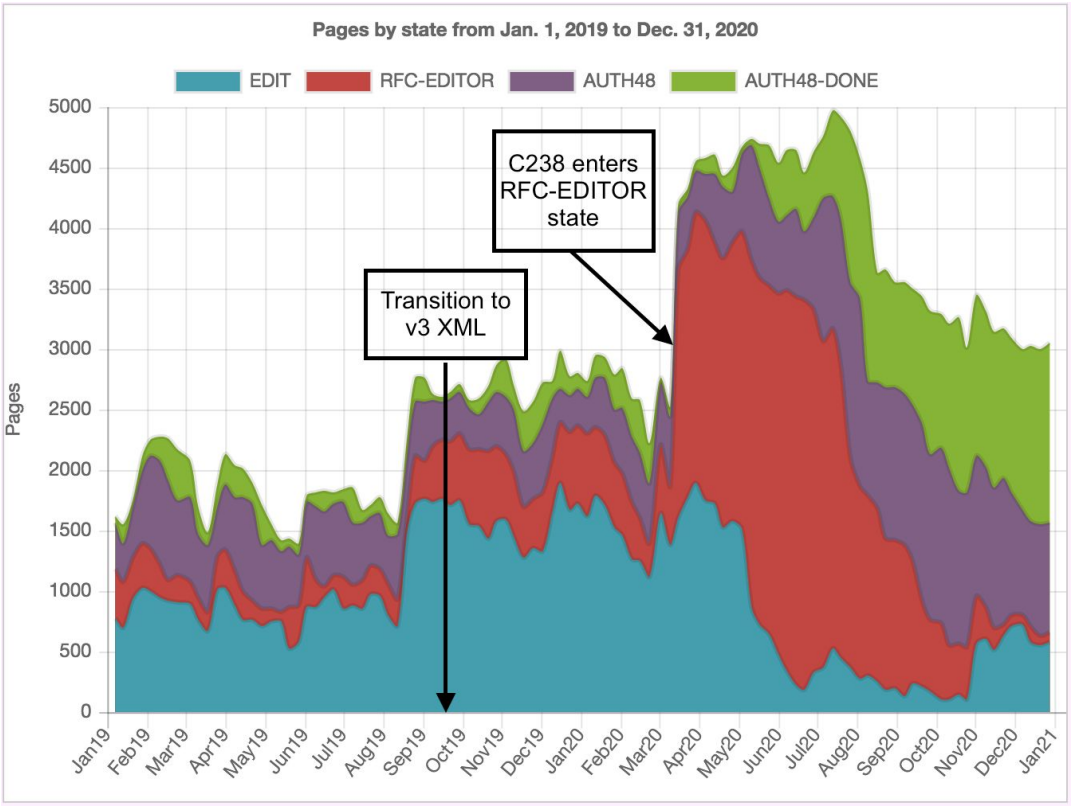

Figure 2

Much of the year was dedicated to converting Internet-Drafts approved for publication from v2 to v3 XML files. Figure 3 shows the type and percentage of source files received and published in 2019 and 2020. Note that only a portion of the documents published in 2019 were published in v3 XML. For I-Ds in which no source was received, text was converted to XML using id2xml.

		v2 XML	None (id2xml)	v3 XML	Total
2019 publications	source received	163	17	0	180
	source used for RFC published	140	N/A	40 (29%)	180
2020 publications	source received	171	13	25	209
	source used for RFC published	0	N/A	209 (100%)	209

Figure 3

Converting files from v2 to v3 XML wasn't as simple as running the v2 XML through the conversion tool. The groundwork in 2019 was to closely study and train on the changes to the vocabulary from v2 XML, on the new features made available with v3 XML, and how to use the new tools. It required experimentation to understand the effects of various tagging and tagging combinations, a review of each of the three output files for display issues, bug reporting, adaptation to the changing toolset and vocabulary, identification of best practices for correct tagging, a review of and updates to internal procedures, and updates to the v3 FAQ (https://www.rfc-editor.org/materials/FAQ-xml2rfcv3.html).

In particular, the RPC marked approximately 15% of the edited docs in 2020 as requiring heavy formatting work (tracking for this began in more detail mid-2020). The following aspects of formatting work were commonly listed as difficult:

- Conversion of <artwork> to <table>
- Difficult table adjustments (e.g., use of rowspan / colspan)
- Conversion of <artwork> to <list>
- Conversion of <artwork> to <sourcecode> and appropriate settings for type=""
- <list>s in general required a lot of review to get the right output
- Cleanup of odd tagging resulting from the conversion tool
- <xref>s - some hardcoded references to sections, document titles, RFC numbers, and combinations thereof are not always straightforward to update
- Line limitations - breaking <sourcecode> lines that extend beyond the 69 character limit (or asking the authors to break them)

In addition, 17% were flagged as requiring more-than-normal cleanup of the content (e.g., syntax, terminology). A given document may have required heavy formatting work as well as extensive content updates.

The transition was made more difficult by the loss of the RFC Series Editor (RSE); the person in that role had evolved the v3 format work from its inception to initial production and was ultimately the community leader who could defend and advocate for the progression of v3 when the processing times became uncomfortable. The RPC welcomed John Levine as the Temporary RFC Series Project Manager in January 2020. John adapted to his role quickly, as he was already familiar with the IETF and RFC Series. However, this role had a narrower scope than the previous RSE. As a result, it seemed there was a lack in gathering community feedback, discerning where action was needed, and making decisions about changes to the v3 XML vocabulary. To help fill this gap, the "RFC XML and Style Guide change management team" was introduced in October 2020.

Until then, the RPC was interacting with the tool developers and members of the tools team frequently to discuss issues as they were discovered. Once the "RFC XML and Style Guide change management team" was formed, the RPC participated in discussion with this team. Many of the issues can be seen as bug reports <https://trac.tools.ietf.org/tools/xml2rfc/trac/report/15>, and the results of many of the discussions are reflected in the xml2rfc FAQ <https://www.rfc-editor.org/materials/FAQ-xml2rfcv3.html> as well as internal documentation.

In Figure 4, the SLA shows that processing times were high throughout much of 2020. However, by year end, many issues related to v3 XML were resolved and the tools were becoming more stable, each of the docs comprising C238 were in AUTH48 or AUTH48-DONE states, and we were on the verge of returning to publishing RFCs in a timely manner. Figure 2 shows the sizable decrease in the number of pages in EDIT and RFC-EDITOR states, and Figure 1 shows a sizable rise in publications in Q3 of 2020. This cleared some of the documents with extended processing times out of the queue. As we see in Figure 4, 17 of the RFCs published in Q4 of 2020 (33%) had processing times of 6 weeks or less. This was a significant improvement compared with the previous 4 quarters. The authors were noticing the decrease in processing times as well, as Carsten Bormann included "We are currently impressed by RFC editor velocity" in the Concise Binary Object Representation Maintenance and Extensions Working Group slides during IETF 109 in November 2020 (see https://datatracker.ietf.org/meeting/109/materials/slides-109-cbor-carstens-slides-00). As of 31 December 2020, the average processing time was 5.9 weeks.

	2019				2020				2019 totals	2020 totals
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4		
Submissions										
Docs	60	53	89	49	65	41	34	36	251	176
PGTE	2057	1491	2296	1576	2537	904	814	1134	7420	5389
Publications										
Pages	1404	1757	1113	878	1383	890	1825	1114	5152	5212
Docs	42	67	31	40	70	34	53	52	180	209
SLA										
Tier 1	7	15	14	0	0	4	0	17	36	21
Tier 2	33	34	17	19	7	5	12	10	103	34
Tier 2 +	2	0	0	21	63	25	41	25	23	154

Figure 4

Below we have provided some feedback we received directly from the authors regarding the quality of the edits and/or the process:

[Redacted]

Challenges

As mentioned earlier, a major challenge for the RPC throughout 2020 was C238 - a cluster of 50 documents that needed to be published together because of normative reference issues. In addition, based on discussion with some authors, our understanding was that changes to text in one document could affect the text in another document. Per a suggestion from Sean Turner, we created the C238 mailing list <https://www.rfc-editor.org/pipermail/c238/> in the hopes of facilitating AUTH48. C238 was released to AUTH48 in batches starting in April 2020; AUTH48 continued into January 2021. Coordinating the changes across so many documents, authors, and editors, as well as continually checking and rechecking portions of a document for outdated material (e.g., outdated references), was extremely difficult. While we hope to never see a cluster of this magnitude again, the RPC will be meeting to discuss possible improvements for cluster handling.

Separately from the already discussed challenges presented in 2020, issues identified in the PDF files of RFCs were addressed in coordination with the Interim RFC Series Project Manager. Specifically:

- The files were not PDF/A-3 as intended (for archival purposes). We corrected the coding error, replaced existing files with PDF/A-3, and advised the community. This did not change the appearance or content of the PDFs.
- WeasyPrint leads to various oddities in the PDF output, primarily regarding vertical white space and the placement of page breaks. There has been ongoing effort to identify and report these items, as well as gain understanding as whether there is a fix available.

Other Updates

In addition to tackling the workload associated with the transition to v3 XML, the RPC made progress in other areas and responded to external requests.

Programming

- Based on the outcome of the security review of the RPC code base (by ZX Security), the programmer made updates to over 50 PHP files used for external and internal web interfaces. After testing, she coordinated with Robert Sparks (IETF Tools Project Manager) to make the code base available in a public repository (https://github.com/rfc-editor/rpcwebsite) in December.

- Links to the HTML files with inline errata were added to the RFC info pages. For example, see <https://www.rfc-editor.org/info/rfc4782> and <https://www.rfc-editor.org/info/rfc5234>.

Process improvements

- The switch to publishing v3 XML and three output files added complexity to our editorial and publication process. To reduce the chance of human error, and to simplify the process, scripts were updated to combine separate steps and archive files appropriately.

- To reduce manual data entry, new scripts were written for pulling an RFC's metadata directly from the XML file for insertion into the database.

- To improve mail handling and archiving, mailman was adopted for <rfc-editor@rfc-editor.org>, an internal list. This was a significant process change for some users.

- To facilitate AUTH48 reviews, a configuration in EMACS was changed based on feedback from an author and the IETF Executive Director.

- To better understand where bottlenecks and challenges in the editing process exist, we improved the tracking of data about the workload associated with each document (from the editor's perspective). More detail (e.g., format work vs. content work) allows better understanding of where editor time is focused and potentially provides suggestions for tool development/improvement.

To fulfill an author request and gain additional experience with GitHub, the RPC used GitHub to interact with the authors and update RFC 8829 during AUTH48. This was quite an undertaking due to the requirement of a new workflow and the number of issues raised. The issues can be

viewed at <https://github.com/rtcweb-wg/jsep/issues?q=is%3Aissue+is%3Aclosed>. More info may also be found at <https://github.com/rtcweb-wg/jsep/issues/1005>.

In addition, the RPC sought volunteers to update the flow chart of the RFC Editor process. Based on input from the RFC Editor, Bob Hinden created the chart available here: https://www.rfc-editor.org/about/queue/flowchart/. Further updates were made to the page

based on feedback from John Klensin.

What's on the Horizon

A primary focus of the RPC is to continue publishing high-quality RFCs, and improve turnaround times in 2021. In addition, we will actively be looking at ways to improve process efficiency and transparency, as well as prepare for potential changes to the RFC Editor Model. Some specific items are discussed below, but we look forward to additional input from the data John Levine has been collecting from the post-publication surveys, as well as input from community leadership, so we can adjust as needed to meet the needs of the community.

AUTH48: We will also be exploring ways to revamp the AUTH48 process, as it's clear from feedback that the current system is less than ideal. Based on more recent discussion, we will consider the requirements for the AUTH48 process (e.g., archival record of author approvals) and the services that might facilitate the process.

Potential experiments: In addition to the evaluation criteria already defined for the "Using GitHub for AUTH48" experiment, we will review the overall objective when considering whether further experimentation is needed. This experiment has been run twice (with RFC 8446 in 2018 and RFC 8829 in 2020-2021) with the evaluation criteria as listed on https://www.rfc-editor.org/rse/wiki/doku.php?id=github_auth48_experiment.

As previously mentioned in other settings, we're also considering whether an experiment using markdown is beneficial. This would mean collaborating with the authors through AUTH48 using markdown and converting the file to v3 XML as a final step in the editorial process.

XML-SG team: We will continue to work with the "RFC XML and Style Guide change management team" to fulfill the goals of:

- Solidify the v3 vocabulary, documenting the changes from RFC 7991
- Define a subset of required semantic tags required for archival v3 XML

Process improvements: We're working with the tools team to push notifications to the datatracker when AD approval is required during AUTH48. This will allow the ADs to more easily track when their approvals are needed. We will also continue to look for ways to refine the editorial and publication process for efficiency. In particular, we will be investigating whether further automation is possible.
We will also be investigating the best way to promote the existence of <https://www.rfc-editor.org/rfc/prerelease/> for authors of bis documents.

Transparency: We will reinstate the RFC Editor meetings with the stream managers to facilitate discussion of stream-wide issues and to ensure that the goals of the streams and RFC Editor are clear and aligned.

Rfced-future: We will continue to track discussion regarding the future of the RFC Series [Editor] and RFC Editor Model to understand the implications of points of consensus and prepare for future implementation.

AMS and the RPC staff are dedicated to providing the Internet Community with first-rate editorial and publication services as well as excellent customer service. We are looking forward to exploring potential process advancements and engaging with the community in 2021.