

RFC 7328 : Writing I-Ds and RFCs using Pandoc and a bit of XML

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 13 août 2014

Date de publication du RFC : Août 2014

<https://www.bortzmeyer.org/7328.html>

Vous voulez écrire un RFC ? Il faut commencer par faire un *"Internet-Draft"* (I-D) et il existe pour cela deux formats, en XML ou bien en Word. Ce RFC présente une troisième solution, utilisant Markdown, ou plutôt sa variante définie par Pandoc.

Le format d'écriture de RFC le plus répandu est sans doute XML, tel que décrit dans le RFC 7749¹. Le format Word est dans le RFC 5385. (Il existe aussi le format historique, utilisant nroff mais je ne suis pas sûr qu'il soit encore accepté aujourd'hui.) Certaines personnes trouvent que les *"tags"* XML les distraient pendant qu'ils écrivent leur *"Internet-Draft"* (I-D) et ils préfèrent donc utiliser un autre format « texte », avec les marquages de Markdown, jugés moins intrusifs. Markdown est très populaire et est utilisé à plein d'endroits différents (par exemple, dans les formulaires du système de suivi de bogues de GitHub). Markdown a plein de dialectes et de variantes et celle utilisée ici est celle du logiciel Pandoc, un sur-ensemble de Markdown. (Par abus de langage, j'écris ensuite « Pandoc » pour désigner ce format.)

(Au passage, vous pouvez regarder la page officielle de Markdown <<http://daringfireball.net/projects/markdown/>> et celle de Pandoc <<http://johnmacfarlane.net/pandoc>>.)

Le document en Pandoc peut ensuite être traduit dans d'autres formats, HTML, EPUB, DocBook, etc, d'où on peut ensuite tirer le futur RFC dans un format de publication officiel (comme le texte seul). L'outil Pandoc2rfc <<http://github.com/miekg/pandoc2rfc>> convertit le Pandoc en Docbook, qui est ensuite traduit en XML-RFC7749, puis ensuite dans les différents formats de publication souhaités (on passe donc d'un texte quasi-brut avec instructions de formatage, à XML, pour revenir à du texte brut...) Dans le futur, peut-être existera-t-il un chemin plus rapide, allant directement du Pandoc au RFC final. En attendant, ce chemin impose quelques contraintes, notamment qu'on ne peut pas complètement éviter d'éditer du XML. L'auteur doit éditer quatre documents :

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc7749.txt>

- Un document Markdown `abstract.mkd` contenant le résumé,
- Un document Markdown `middle.mkd` contenant le gros du RFC,
- Un document Markdown facultatif `back.mkd` contenant les annexes,
- Un document XML `template.xml` contenant auteurs et références bibliographiques.

Un exemple du gabarit `template.xml` figure dans la section 2 de notre RFC.

Pour installer l'outil `Pandoc2rfc` <<http://github.com/miekg/pandoc2rfc>>, vous allez avoir besoin du processeur XSLT `xsltproc` (la transformation du Docbook en XML-RFC7749 est faite avec XSLT) et bien sûr de Pandoc <<http://johnmacfarlane.net/pandoc/>> lui-même.

`Pandoc2rfc` contient un script shell qui fait les transformations nécessaires (voir la section 3 du RFC).

Pour apprendre la syntaxe de Pandoc, il faut regarder le guide <<http://johnmacfarlane.net/pandoc/README.html>>. `Pandoc2rfc` traduit via `xml2rfc` (l'outil du RFC 7749) et met en correspondance les concepts Markdown avec ceux de XML-RFC7749. Par exemple, une énumération avec `*` va être traduite en une liste (<`list`> XML) `style="symbols"` et une énumération avec des numéros (comme "1.", "2.", etc) va être traduite en `style="numbers"` (RFC 7749, section 2.22.3). La section 4 de notre RFC donne une liste complète des équivalences Pandoc;-XML, résumées sous forme d'un tableau dans l'annexe B.

Il est d'ailleurs difficile de faire du `Pandoc2rfc` sans connaître `xml2rfc`. Par exemple, pour afficher une table des matières, il faut inclure l'invocation `xml2rfc <?rfc toc="yes"?>`.

La section 5 liste quelques limitations de `Pandoc2rfc` par rapport à Pandoc : pas de gestion des index, les commentaires doivent être faits avec des commentaires HTML dans le source Pandoc, les citations se font avec les références XML, le système de citation de Pandoc n'est pas utilisé, et enfin les auteurs doivent se résigner à voir apparaître des erreurs de `xml2rfc`, la traduction Pandoc vers Docbook ne gère pas toutes les erreurs possibles. De mon expérience, il faut pas mal bricoler pour faire marcher le tout. Écrire des RFC en Markdown est sympa mais on n'y arrive pas en cinq minutes.

Si vous voulez un exemple concret, la distribution de `Pandoc2rfc` <<https://github.com/miekg/pandoc2rfc>> contient un répertoire `test/` avec un exemple de RFC bidon utilisant toutes les fonctions de `Pandoc2rfc`. J'ai un peu modifié le `Makefile` pour faire des versions texte, XML et HTML du RFC :

```
% make all
pandoc2rfc -t template.xml back.mkd middle.mkd
pandoc2rfc -H -t template.xml back.mkd middle.mkd
pandoc2rfc -X -t template.xml back.mkd middle.mkd
```

Et cela produit les trois versions, à partir du source Markdown (essentiellement en `middle.mkd`) et du gabarit `template.xml`.