

# Experiences on an Open Source Translation Effort in Japan

Hiroki SATO\* Keitaro SEKINE  
*hrs@jp.FreeBSD.org*

Faculty of Science and Technology, Tokyo University of Science, JAPAN  
FreeBSD Japanese Documentation Project

BSDCon 2002, February 13, San Francisco, California, USA

# Agenda

1. Introduction of FreeBSD Japanese Documentation Project
2. What is translation work?
3. Problems and what we need
4. Effective ways for the work
5. Summary

# Introduction of FreeBSD doc-jp

## The FreeBSD Japanese Documentation Project

- started by some FreeBSD developers in Japan in 1996.
- main goal: translation of FreeBSD-related documents from FreeBSD Documentation Project(FDP) such as FreeBSD Handbook.
- two leading efforts: Japanese Manual Project(JPMAN) and doc-jp.

# Results of doc-jp (1)

- FreeBSD Handbook (1996-, printed version was also published in 1997)
- FreeBSD FAQ (1997-)
- [www.FreeBSD.org](http://www.FreeBSD.org) (1997-)
- FreeBSD Release Notes/Errata (1997-, per release basis)
- FreeBSD Security Advisory (2000-, per release basis)

## Results of doc-jp (2)

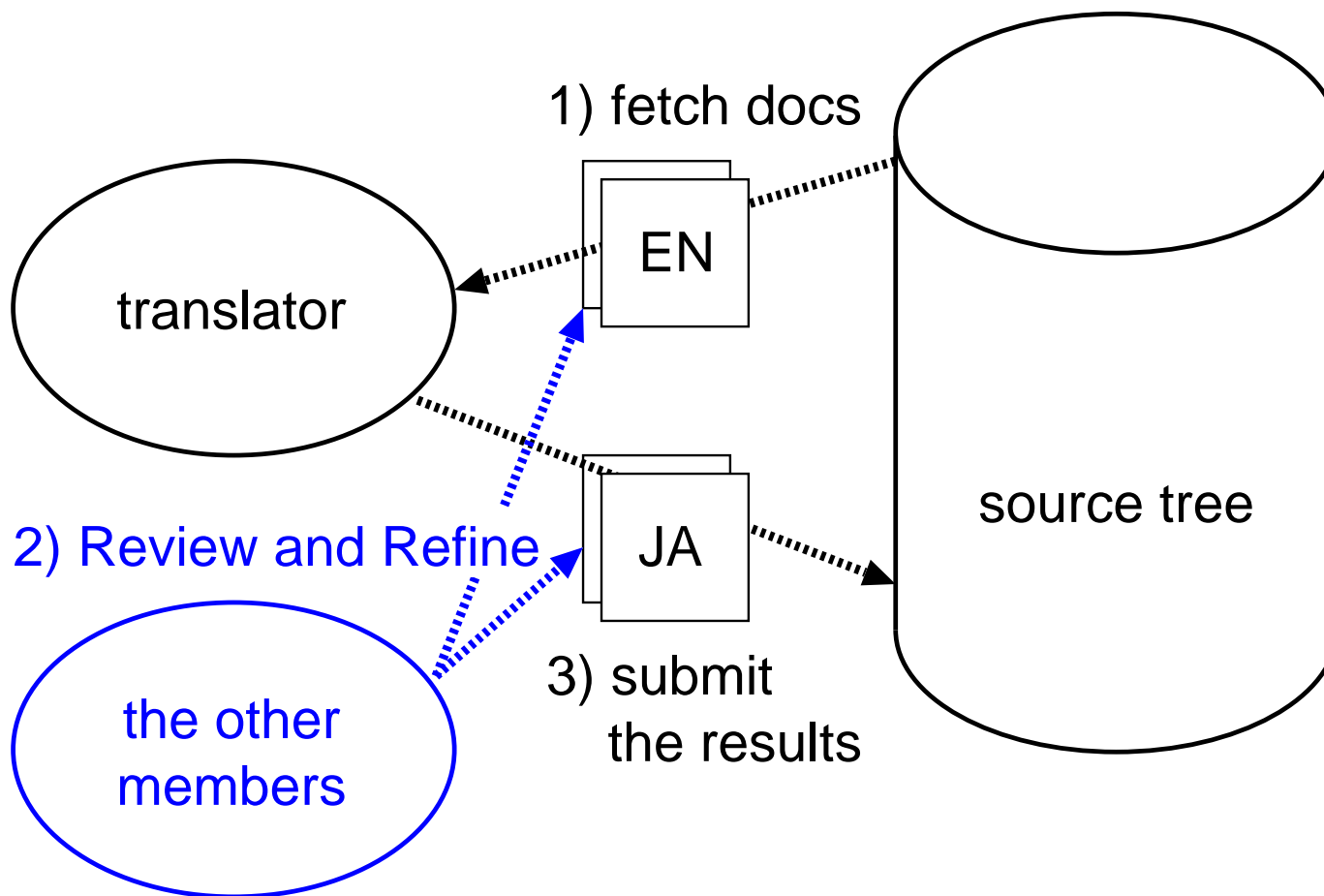
The numbers of original and translated documents

location of the source tree	en	ja	%
doc/\${LANG}/articles	26	8	30.8%
doc/\${LANG}/books	72	47	65.3%
www/\${LANG}	203	153	75.4%
src/release/doc/\${LANG}	22	7	31.8%
<b>Total</b>	<b>323</b>	<b>215</b>	<b>66.6%</b>

Note: “en” in this table includes documents which we could not translate due to copyright issue.

# What is translation work?

## Typical Translation Flow



# Peculiarities (1)

## Quality Evaluation

code development

- determined by its behavior itself.
- It is almost objective.

document translation

- needs reviews of the readers.
- It is basically subjective.

# Peculiarities (2)

## Life Span of Translated Documents

### code development

- Normally it is updated to add new features or to fix bugs.
- We can usually use the old one if no fatal bug exists.

### document translation

- Normally it is updated when it becomes obsolete.
- We cannot use the older translated documents.

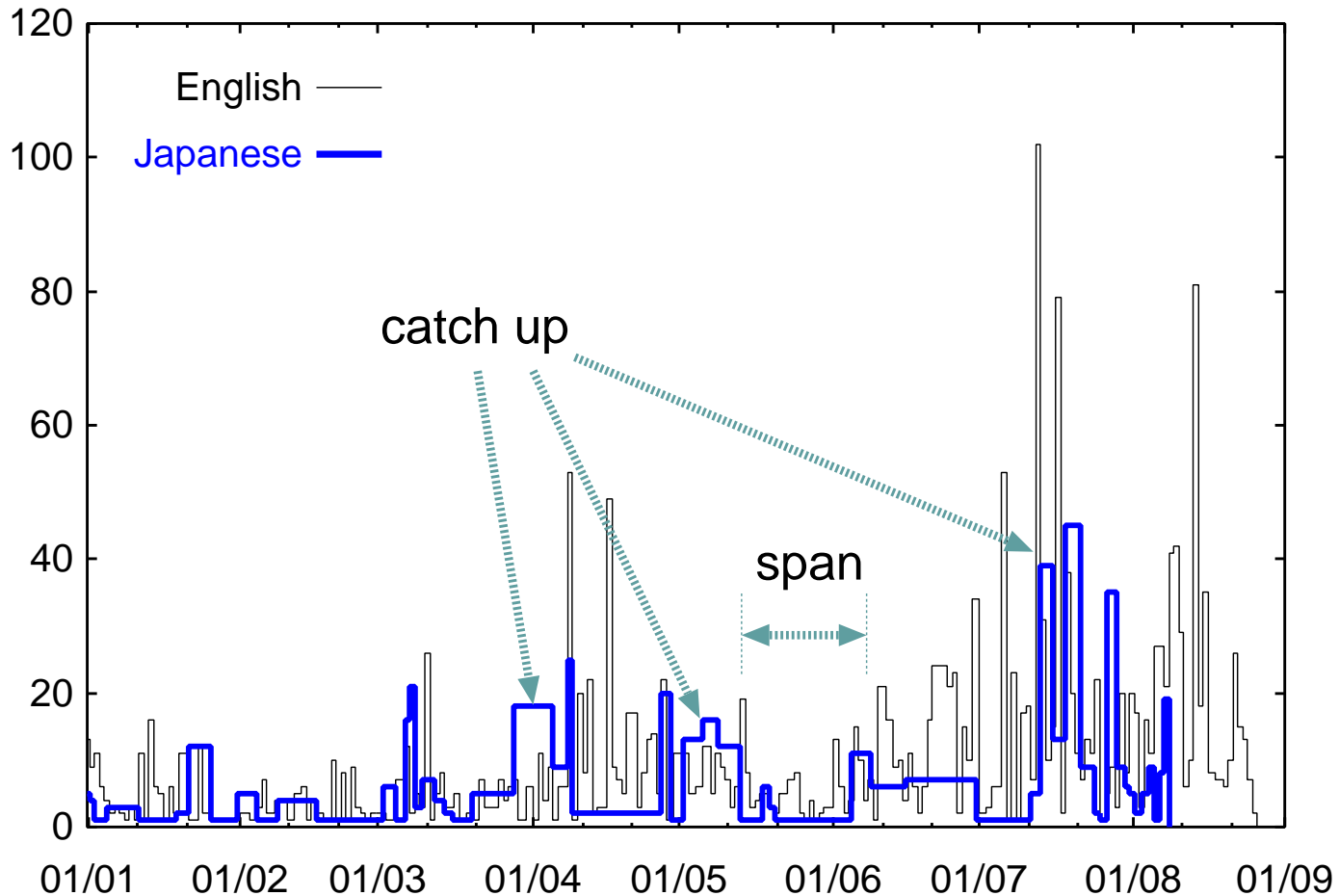


- The quality evaluation is very difficult.
- The translated document becomes useless (and sometimes harmful for readers) when the original document is revised.

⇒ continuous effort is required

- We need the process “review and refine” again and again.
- To keep the project members motivated high is also essential.

# Troubles(1): Release Engineering



commit frequency to FreeBSD doc/ and www/ tree

# Troubles(1): Release Engineering

The translation work is always behind.



So it is difficult to release the documents like release notes or security advisory.



**We need the time to translate**

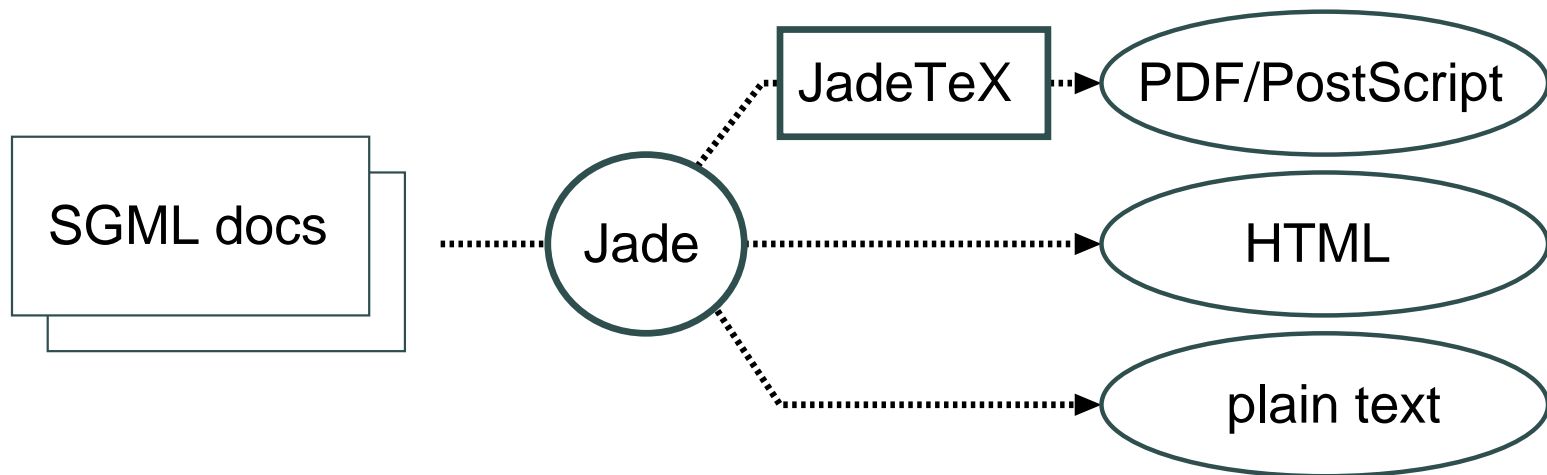
# Troubles(1): Release Engineering

- For translation teams, the period “documentation freeze” would greatly help.
- More translators’ efforts are required, of course :-P

Incidentally, at present the first draft of FreeBSD-SA Japanese translation is available within 24 hours from the time the advisory is published.

# Troubles(2): Toolchain

## Processing SGML Documents



- "Jade" processes SGML docs using DSSSL stylesheet.
- "JadeTeX" is a back-end for PDF/PS output.

**At present, JadeTeX cannot handle Japanese characters.**

# Troubles(3): separate commits

## the rule of separate commits

- “cosmetic changes”: changing locations of line-breaks, fixing typos, and so on.  
⇒ We can follow the changes without translation skill.
- “content changes”: rewriting sentences themselves.  
⇒ These need translation skill.

# Troubles(3): separate commits

## In case of the rule is not followed

Before:

This is a **simple** document.  
The rule of "separate  
commits" originated in the  
FDP for the translation teams.

revised!

After:

This is a **sample** document.  
The rule of "separate commits"  
originated in the FDP for the  
translation teams.

The diff delta:

- This is a **simple** document.  
- The rule of "separate  
commits" originated in the  
- FDP for the translation teams.  
+ This is a **sample** document.  
+ The rule of "separare commits"  
+ originated in the FDP for the  
+ translation teams.

we should carefully compare the large deltas.

# Troubles(3): separate commits

## In case of the rule is followed

Before:

This is a **simple** document.  
The rule of "separate  
commits" originated in the  
FDP for the translation teams.

revised!

After:

This is a **sample** document.  
The rule of "separate commits"  
originated in the FDP for the  
translation teams.

The separated delta:

- This is a **simple** document.  
+ This is a **sample** document.

- The rule of "separate  
- commits" originated in the  
- FDP for the translation teams.  
+ The rule of "separare commits"  
+ originated in the FDP for the  
+ translation teams.

with CVS log:  
"cosmetic changes only"

we can easily understand where is changed.



# Troubles(3): separate commits

## Interchanging Lines

Before:

This is a simple document.  
The rule of "separate commits"  
originated in the FDP for the  
translation teams.

revised!

After:

The rule of "separate commits"  
originated in the FDP for the  
translation teams.  
This is a simple document.

The diff delta:

```
- This is a simple document.  
The rule of "separated commits"  
originated in the FDP for the  
translation teams.  
+ This is a simple document.
```

the same sentence is  
deleted and added

this will confuse the translator.

# For Efficiency(1)

## Providing Text fragments

- Most of the project members do not know the latest status of the original documents.
- People who want to join us often hesitate; clarify what should be translated and how to do so.
- To split original documents and provide them into the project's mailing list is effective.

# For Efficiency(2)

## Notification of the revision

original document:

```
$FreeBSD: doc/.../book.sgml,v 1.119 2001/11/19 ...$
```

translated document:

```
Original revision: 1.118  
$FreeBSD: doc/.../book.sgml,v 1.70 2001/11/19 ....$
```

compare

CVS ID

insert when processing

Note: This document is obsolete!

# For Efficiency(3)

## Translation table

- useful to translate and keep consistency.
- word-by-word style: it is used widely.
- sentence-by-sentence style: it can reflect the context and can be generated from the translated documents.

# Summary

- The translation work is essential in the countries like Japan. This is because most of the people cannot speak English.
- The work is not temporary one: it needs continuous effort to keep the quality high and the documents up-to-date.
- The translators need some helps: “documentation freeze” and “the rule of separate commits.”

# Summary *(continued)*

- The several ways for efficient translation work are proposed.
  1. providing split documents actively to the project members is effective to keep them motivated high.
  2. notifying the delay of translation.
  3. maintaining translation table is also useful.