Synchronization problems in parallel activity of SmartScribo blog processors

Rustam V. Kadirov, Dmitry G. Korzun, Ivan V. Galov

Petrozavodsk State University
Department of Computer Science

This project is supported by grant KA179 of Karelia ENPI - joint program of the European Union, Russian Federation and the Republic of Finland

11th FRUCT conference
April 23–27, Saint-Petersburg, Russia
Table of Contents

1 Review of Smart-M3 platform
2 Infrastructure of blog processors
3 The task of synchronizing blog processors
4 The task of authorization on services
5 The task of data integrity
6 Conclusion
SmartScribo system

Problems:

- RSS and Twitter in SmartScribo system
- Blog processors classification
- Authorization on blog services
- Protocol for synchronization interaction of blog processors
- Problem of data consistency
Smart-M3 platform

- Open source software project
- For each service own agent
- Ontological representation of data
- SmartSlog as Smart-M3 SDK
SmartScribo blog processors

- Registration/authorization of user
- Receiving information about user
- Operate with posts and comments
- Operate with friends
- Multiblogging
- Notifications system
## Classification of blog processors

<table>
<thead>
<tr>
<th>Type</th>
<th>Subtype</th>
<th>Blog processor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bidirectional</td>
<td>Standart</td>
<td>BP-LiveJournal</td>
</tr>
<tr>
<td></td>
<td>Reducing</td>
<td>BP-Twitter</td>
</tr>
<tr>
<td></td>
<td>Mixing</td>
<td>None</td>
</tr>
<tr>
<td>Only reading</td>
<td>With registration</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Without registration</td>
<td>BP-RSS</td>
</tr>
<tr>
<td>Multiblog</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Integration</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>
RSS in SmartScribo

- Describing rss feeds and changes in blogs
- Without registration
- Access only for reading
Current implementation (BP-RSS)

- Work in read-only mode with following blog services: Twitter, LiveJournal, WordPress Digg

- RSS support was implemented in following blog clients: Windows Maemo5
Blog service Twitter

Twitter is reduced blog service

- Length of messages less than 140 symbols
- Messages do not have comments
- System of friends is replaced by system of followers
Infrastructure of blog processors

Currents implementation (BP-TW)

- Authorization on service
- Send or delete post on service
- Receiving list of posts
- Using of SmartSlog library
- Blog processor Twitter support in Maemo5 and Windows blog clients

Rustam Kadirov
The task of synchronizing blog processors

Duplication of information with two or more blog processors

- Two or more blog processors operating with same blog services
- One notification from any blog client
- Duplication of information in Smart Space blogosphere
Synchronization of work

- Modification of notification system
  - Sending current state of blog processors
  - Receiving current state of blog processors on blog client side
  - Sending notification to first free blog processor

- Reducing duplication of knowledge

Blogosphere Smart Space

- BP-TW-1, serviceName, Twitter
- BP-TW-1, notificationPrefix, TW
- BP-TW-1, state, idle

- BP-TW-2, serviceName, Twitter
- BP-TW-2, notificationPrefix, TW
- BP-TW-2, state, idle

- Notification: BP-TW-1, refreshPosts, account-1
- account-1, login, user-1
- account-1, password, password-1
- account-1, hasPost, post-1
- post-1, title, Some Title
- post-1, text, Hello SmartScribo!

Blog Service

- BLOG PROCESSOR-TWITTER-1
  - 1 Sending state
  - 4 Getting notification
  - 5 Getting new posts
  - 6 Receiving new posts

- BLOG PROCESSOR-TWITTER-2
  - 7 Publishing of new posts and change state

8 Receiving new posts

2 Extracting state of BP

3 Sending notification
The task of authorization on services

Store personal information needed for authorization

- Problem of storage with oauth keys and password of user
- Threat of security

Possible solution:
- Waiting new version of Smart-M3 with some security system
- Own system of security
The task of data integrity

Dynamical join/leave of blog processor

- Lost information because of unexpected appearance and disappearance of blog processors
- Losing of user data or some links between data

Special knowledge processor for analysing of Blogosphere Smart Spaces
  - Removing duplication of data
  - Recovery of lost data
Metrics of project

■ BP-RSS:
  ▶ Python 2.7.1+
  ▶ M3-Python KPI 0.9.2 library
  ▶ Python feed parser 5.1.1 library

■ BP-Twitter
  ▶ ANSI C 99
  ▶ SmartSlog SDK 0.37alpha
  ▶ OAuth 0.9.6 library
  ▶ JSON-C-parser 0.9

<table>
<thead>
<tr>
<th>Blog processor</th>
<th>Code lines</th>
<th>Comment lines</th>
<th>Total lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP-RSS</td>
<td>680</td>
<td>126</td>
<td>970</td>
</tr>
<tr>
<td>BP-Twitter</td>
<td>2285</td>
<td>838</td>
<td>3123</td>
</tr>
</tbody>
</table>
Results

- Classification of blog services
- Blog processor-RSS
- Blog processor-Twitter
- All blog processors deployed on server maemo-m3.cs.karelia.ru
- Realization of authorization on services, which use OAuth
- Protocol of synchronization interaction of blog processors for reducing duplication of data
- Concern the problem of data integrity