Codeathon round-up

Brion Vibber - Wikimedia Foundation
Wikimanía - Buenos Aires - 28 de agosto 2009
http://techblog.wikimedia.org/
We had a little less structure for the first planning day than we did at the April developer meeting; next time maybe target somewhere in the middle. Having the room available throughout the conference was very valuable, providing a workspace and meeting space for techies as well as a quieter place to hang out for folks to get a break from the busier rooms upstairs.
¡Muchas gracias!

Facilities came together great .... it’s not a Wikimania without wifi failures, but it works more than it doesn’t work and they’ve fed us well ;)

We’ve caught up on back review to the last branch point in June and have cleaned up most of the blocking issues -- expect updated code before you get home. :)

Lots of big changes coming internally:
* js2 branch merge
* major maintenance scripts cleanup
* redone upload backend (supports background fetches of large files, pretty new interface)
* new localization caching framework (better support for LocalisationUpdate)
Maps!

Got the OSM integration dev team talking with the ops team; we should have the server situation sorted out shortly and be ready for test deployments of SlippyMap soon. Toolserver experimentation has been going for some time with good results; once the first set of official servers are happy we’ve got additional servers for the US data centers budgeted.
Configuration database!

Worked out some basic specs for a configuration database, which will let us migrate to a cleaner way of handling single-wiki and multi-wiki configurations. More minor updates will be possible for admins and stewards on the wiki, reducing bottlenecks where Wikimedia ops techs are required to tweak things.

Migration and backwards-compatibility for old settings are definitely a concern: we'll want to make sure that when a setting is overridden by a LocalSettings.php variable that this is indicated clearly in the config UI.
Good talk w/ Domas, Mark & Tim about profiling and monitoring infrastructure and some things we can plug in -- integrating profiling into PHP, building an aggregator of currently running requests and functions, better detection of PHP failures.

<table>
<thead>
<tr>
<th>Status</th>
<th>Last Check</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>WARNING</td>
<td>08-28-2009 16:09:04</td>
<td>1d 0h 56m 8s</td>
</tr>
<tr>
<td>WARNING</td>
<td>08-28-2009 16:09:02</td>
<td>2d 4h 6m 38s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:10</td>
<td>4d 18h 35m 44s</td>
</tr>
<tr>
<td>WARNING</td>
<td>08-28-2009 16:09:04</td>
<td>5d 18h 2m 47s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:10</td>
<td>6d 3h 35m 47s</td>
</tr>
<tr>
<td>WARNING</td>
<td>08-28-2009 16:09:03</td>
<td>16d 15h 16m 39s</td>
</tr>
<tr>
<td>WARNING</td>
<td>08-28-2009 16:09:08</td>
<td>14d 13h 39m 27s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:14</td>
<td>65d 21h 20m 2s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:12</td>
<td>65d 18h 10m 22s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:14</td>
<td>65d 18h 11m 2s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:02</td>
<td>65d 21h 19m 59s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:03</td>
<td>65d 21h 19m 39s</td>
</tr>
<tr>
<td>CRITICAL</td>
<td>08-28-2009 16:09:04</td>
<td>45d 21h 51m 29s</td>
</tr>
</tbody>
</table>
Domas complained about the mb_string compatibility functions being incredibly slow and this causing serious problems for some users. mb_substr implementation was insane; now has been replaced. :)

```php
function mb_substr( $str, $start ) {
    $ar = array();
    preg_match_all( '/./us', $str, $ar );
    if( func_num_args() >= 3 ) {
        $end = func_get_arg( 2 );
        return join( '', array_slice( $ar[0], $start, $end ) );
    } else {
        return join( '', array_slice( $ar[0], $start ) );
    }
}
```
More and more tools coming into TranslateWiki -- now supports localizations for XUL apps thanks to the okawix team. One current issue is performance on completion status reports -- needs to be made faster so localization teams can organize more actively.
LocalisationUpdate update

I received this e-mail from Roan Kattouw with some great news .... He is part of the Usability Initiative team..

"Just now I've enabled the LocalisationUpdate on the prototype sandbox so we have a staging area for it. Sandbox is running r54813 (others: please don't update the core code on sandbox for a while, kthx). You can see that a bunch of Afrikaans localizations for the toolbar was added in r54979 [1] but are visible on sandbox [2], even though it's running r54813 [3]. The toolbar on Wikipedia is not translated to Afrikaans at all right now [4].

Roan Kattouw (Catrope)"

Actually enabled on test.wikipedia but doesn’t quite work yet. ;) Should be ready to roll when the code push comes in a couple days.
Roan set up a temporary upload space for the Wikimania videos with the awesome local team have recorded and encoded as Ogg Theora for us, yay! We should have them transferred to Commons soon. (Michael Dale’s upcoming improvements to large uploads will make this easier shortly, though we’re still looking at the possibilities for sftp batch uploads too.)
Test site up with planned enwiki configuration -- we hope to push into limited usage within a few weeks. Also being used in similar configuration on MediaWiki site for manual and extension pages, which is helping to get more developer buy-in.
Andrew has been doing amazing things fixing up the old LiquidThreads project providing a cleaner way to handle threaded discussion pages. Now working on the front end, making it AWESOME and inviting to use.
Andrew and Roan have also been poking at other fun things, including a very cool experiment with AJAX section editing. New work is using jQuery as base library, to much happiness. Many related tools to examine and integrate or adapt.
Semantic MediaWiki testing?

Chatted a little with Markus & Danny -- we’re talking about the possibility of setting up an official test using a shared SMW data store. See what the performance and workflow really looks like at Wikipedia scale!
WikiTrust testing?

Planning to get this going soon (unfortunately just haven’t had a chance to run into Luca yet this week ;)

Offline readers

Tomasz – work on OpenZIM common format & reader library, etc. Much interest in using WikiTrust data to aid in selecting good revisions to push out for offline use.
Fixed a bug in Croatian localization. :) Some cool updates to the UI are coming soon...
Domas and Tim have come up with a better way to handle MySQL replication lag monitoring for large batch operations. While not yet implemented, this should be super happy fun time! Now we just need to find a better way to break up the batch operations so we don’t lag in the first place ;)

http://commons.wikimedia.org/wiki/File:The_stopwatch_timer_is_running_in_the_background.PNG
Bandwidth optimization

Mark & Domas are clearing out unnecessary debug headers from our HTTP stream and tuning our compression -- saving about 0.5 megabits/s. Ongoing work in dev trunk to trim down our HTML and JS as well...
On our ops side, Rob is suggesting a coordinated effort to clean up our ops documentation on wikitech.wikimedia.org. I for one welcome our new documentation overlords :D