



OPEN SOURCES #22: THE TORONTO CODE SPRINT

MICHAEL GERLEK REPORTS ON AN OSGEO OPEN SOURCE WEEKEND EVENT TO BE REMEMBERED

Some years ago, the stereotypical computer programmer worked alone in dark basements, fueled only by caffeine and pizza. Thankfully that stereotype has fallen out of favor (except in movies) – unless you are an open source computer programmer, in which case not only are you expected to live in a dank cave and drink Jolt cola but you are also expected to be unshaven, neurotic, and antisocial.

In reality, of course, this is not true, and a recent event in Toronto proved just how social – and productive – open source programmers can be.

The Need for Sprints

At the company I work for, all the developers work in the same office. This means that when something needs to be discussed – determining the best way to implement a new feature, say, or understanding the causes of a hard bug – we can all gather around a white board and hash things out. In the open source world, however, more often than not the development work is done by people from all around the world who have never met each other face to face and only work on the project part time. Indeed, the people on a project may not even share fluency in a common language. Issues are discussed and negotiated at one remove, using email, IM or IRC, and wikis.

By any large, this works well (and better, in fact, than some

would believe). But sometimes the need for in-person collaboration is too strong. This can be the case when significant new features are being designed, when a new release requires a coordinated and dedicated push to completion, or when a backlog of bugs requires a team effort to triage and resolve.

A “sprint” is a software engineering term used to describe a time period during which a specific set of tasks is to be accomplished; the term often connotes an all-hands, heads-down effort. In the open source world, a code sprint (or, in some circles, a “hackathon”) is when a group of programmers gather together, live and in person, to work on their project for a short time. This is often done over a weekend before or after a conference, when the team members are likely to already be in the same city. Attendance is usually open to all comers, but the expectation is typically that you are already a known project contributor and will be able to plunge into the project without any learning curve.

The OSGeo Toronto Sprint

Imagine, if you will, spending a long weekend with some of the brightest minds of the open source geo world, major contributors on everything from GDAL to MapServer to PostGIS to OpenLayers. I had that opportunity in March of this year, when a couple dozen of those people gathered at a hotel in downtown Toronto for a four day OSGeo code sprint [1].

Each morning, we all met in a windowless (but networked) hotel conference room. Each person sat down at one of a number of large round tables, each loosely associated with a particular project, and opened up their laptop and got to work. Heavy discussion – and light banter – flowed fast and furious around the room all day long. Lunch was generally impromptu take-out and refreshments were smuggled in to keep us going. At the end of the day, we folded up our laptops and went out for a group dinner. (And late one afternoon we all quit work a bit early to attend a minor league hockey game, hosted by one of our group who happened to be a former professional hockey player. How cool is that?)

Funding for the event was provided by seven corporate sponsors who each contributed about US\$500. (Disclaimer: my company was one of the sponsors.) These funds covered the cost of the conference room, some snacks, the nightly dinners, and *le hockey*; attendees were responsible only for their own travel and lodging. The final financial report for the sprint, which of course was posted in detail on the web, showed a \$20 surplus (which was donated back to OSGeo).

The Results

So, what did we accomplish during those four days? A number of people blogged about the event in detail [2], but here are some of the highlights:

MapServer: lots of work done, including design work on new XML-based map files, discussions about adding pluggable renderers (OpenGL!), more documentation, and some performance work to improve the speed of EPSG code lookups

PostGIS: an extensive planning exercise for “PostGIS 2.0”, including the “WKT Raster” piece which has received some attention lately on blogs and mailing lists

GDAL: Various bugs, both in the area of functionality and performance, were addressed. In one case I was personally involved with, a few of us in the room were able to work together to collect data and analyze a long-standing performance issue reading MrSID files with the GDAL library. Immodestly quoting from Paul Ramsey’s blog

[3]: The face-to-face communications bandwidth is so much higher that problems fall by the wayside at a great rate. I saw Frank Warmerdam reviewing MrSID code in GDAL with Michael Gerlek [and Chris Schmidt], and Michael being able to quickly point out performance mistakes: “don’t call that function every time”, “these functions are costly, but only the first time through”.

Addressing this bug will improve the speed of applications like MapServer that use GDAL for the MrSID image format and other similar formats.

Contracted work: For a different project I was involved in, my company had contracted with one of the maintainers of a relatively new open source library to make some changes and add some features. Since the principal maintainer of the package was an attendee at the sprint, we were able to discuss the changes he’d made and do some review of the work he’d done in person, rather than by phone or email.

Demos: In a nice piece of marketing and outreach, Chris Schmidt reused his code for OpenLayer’s famous “demo gallery” to become a demo gallery for OSGeo itself [4].

Thanks

The Toronto sprint was by all measures a great success and we expect to have another one early in 2010. Other sprints happen all the time all over the world for a variety of projects. Watch the mailing lists for the projects you’re interested in. If you’re a developer (or UI designer, or technical writer, or…) you might want to consider joining in on the fun. And even if you are “just” a user and can’t contribute directly to the project’s code base, feel free and stop by to let the folks know you appreciate their hard, and often unpaid, work. Finally, I’ll note that while open source projects tend to be very loosely organized, grass-roots efforts, in this case Paul Ramsey of Clever Elephant, along with local organizers Tom Kralidis and Jeff McKenna, did most of the heavy lifting



The Code Sprint was not all work! Participants check out one of Toronto’s fine pubs.



Some of the Toronto Code Sprinters hard at work



It wasn’t all fun and games – some serious work was accomplished

to host the event, including rounding up the sponsors who made it all possible. We thank them and the sponsors for their support of open source in the geospatial community.

Links

- [1] http://wiki.osgeo.org/wiki/Toronto_Code_Sprint_2009
- [2] <http://blog.cleverelephant.ca/2009/03/toronto-code-sprint-recaps.html>
- [3] <http://blog.cleverelephant.ca/2009/03/sprint-day-4.html>
- [4] <http://gallery.osgeo.org>

Author Bio – Michael P. Gerlek (mpg@lizardtech.com) manages the Engineering team at LizardTech and helps develop technologies for managing large-scale GIS datasets. He has been associated with OSGeo since its inception and champions the use of open source software in the commercial sector. Michael lives on an island in the Pacific Northwest.