Looking for Google Code-in? Click here!

Sharing Metadata Between Multimedia Applications Using Baloo

steckdenis

Short description: Baloo, the successor of Nepomuk, is a library allowing applications to associate metadata with files or other resources. This proposal is about using Baloo to store play-lists, statistics and sharing information about multimedia files. This way, such information would be shareable between multimedia applications.

Name: Denis Steckelmacher
Email Address: steckdenis@yahoo.fr
Freenode IRC Nick: steckdenis
IM Service and Username: Skype, steckdenis@yahoo.fr (I very much prefer e-mails or IRC)
Location (City, Country and/or Time Zone): Brussels, Belgium, GMT+1
Proposal Title: Sharing Metadata Between Multimedia Applications Using Baloo

Motivation for Proposal / Goal

Currently, each multimedia application of KDE uses its own implementation of playlists, most played files, tags and score. This idea comes from a dream of Aaron Seigo, who said that it would be wonderful for the users to be able to arrange their multimedia files in collections using Juk or Amarok, and to be able to see their collections in Plasma Active, Plasma Media Center on their TV or in any other multimedia application able to use Baloo.

This feature was already thought of several years ago, but it was when Baloo was still Nepomuk (which was too slow for big collections). Now, Baloo uses a Xapian-based database that is way faster than Nepomuk, according to the Baloo developers and some other ones.

Implementation Details

This proposal is about doing as much as possible during the summer in order to integrate the multimedia applications with Baloo. The goal expressed above it the main target of the project, but several smaller projects are needed or possible to do during the summer to reach the goal:

- Extending Baloo so that it can extract metadata information from even more multimedia files (taglib is already supported, but EXIF and format-specific metadata could be useful.
- Adding to Baloo the ability to write metadata directly into files, and not only in its database. It is a very interesting project technically speaking and would allow the user to associate metadata to their multimedia files (artist, album, title, year, etc) using a Baloo-powered application and having these written back into the files so that non-Baloo applications can see them. This will make this entire project interoperable with non-Baloo applications.
- Building a library that downloads metadata about multimedia files on IMDB or any other online database. The metadata will be written in Baloo (and the files themselves following the point above). This way, all the KDE multimedia applications will not have to duplicate this downloading and parsing code, and even file managers (if they want to) could have a button allowing the user to download metadata about files. Nepomuk had its webextractor library that could be used as a base or simply ported to Baloo.
- Discussing with the multimedia application developers (especially the developers of Amarok) and seeing how best to store information about

playlists, most recently played, most played and preferred songs in Baloo. Some applications may even be modified to use Baloo instead of their own implementation of playlists.
- Modifying KDE applications to make use of the new Baloo features. A simple multimedia application which I use everyday is **Juk**, so I plan to work on this application first, in order for Baloo to support at least all the features of Juk. If there is a better application for these kinds of experiments, I would be glad to work on it.

Tentative Timeline

The community bonding period will be used to discuss with the multimedia application authors about everything they may want to see added to or modified in Baloo. I will also discuss with the Baloo developers to decide which parts of the project will be done in Baloo Core and which ones could be subject to a new library. Baloo also has a repository for metadata extractors and widgets. As this project will touch many parts of Baloo, each new feature must be placed somewhere.

Here is a tentative timeline for the 14 weeks of the GSoC and the two weeks after the project. The good thing is that university starts at the end of September in Belgium, so I will have plenty of time for KDE after the coding period.

1. Adding file extractors for common multimedia file formats not handled by taglib. This will allow me to code something while still thinking about everything discussed during the community bonding period.
2. Adding even more file extractors if needed, and starting work on the infrastructure needed to write Baloo metadata back into files. Whether this is done every time a file metadata is modified or periodically by a server still needs to be discussed
3. contd.
4. End of what has started on week 2. Normally, after 3 weeks, the feature should start to be usable.
5. Seeing if the Nepomuk webextractors library can be ported to Baloo, and beginning the porting process. At this time, I cannot say whether porting webextractors or reimplementing it (using code from it when possible) will yield the best results.
6. contd.
7. Webextractors should be ported by this week, but it lacks an IMDB extractor. This week will be used to add some missing extractors for well-known online databases.
8. Adding a Baloo-based library that can be used to organize multimedia files in playlists. The notion of "playlist" could also be extended a bit, as Aaron Seigo suggested that even non-multimedia files should be associated with lists (a list could for instance contain songs, videos and bookmarked URLs or photos). Every application then only displays the files it supports.
9. Modification of Juk for it to take advantage of Baloo. This will allow me to ensure that everything goes well and that Baloo is up to the task. This point depends on point 8, but these two points can be moved all the way up in the timeline if someone wants to have this checkpoint before the midterm evaluation (it makes sense to first ensure that Baloo works for multimedia applications before extending it for this use-case).
10. Porting Juk to Baloo (continued)
11. This week and the three next ones are a buffer that still needs to be filled. Libraries will need to be documented, bugs to be solved, and Juk is not the only multimedia application of KDE. Plasma Active or Plasma Media Center are possible candidates. Maybe XBMC, a non-KDE set of application, could be modified to take advantage of Baloo. This way, the many users using XBMC as a multimedia station and KDE as their primary desktop will be able to organize their songs on their computer and to get back their playlists on XBMC.
12. contd.
13. contd.
14. contd.
15. After the end of the GSoC, if I didn't have the time to do so before, I would be glad to work on Amarok and to see to what extent Baloo can be integrated into it. Amarok already has its own database, but maybe some sort of bridge between it and Baloo could be implemented. I also still have some things not related to multimedia that I would like to see implemented in Baloo, if I don't have time do to them before the GSoC.

Do you have other obligations from late May to early August (school, work, vacation, etc.)?

The university courses end at the end of May, so I could be unavailable the very first days of the GSoC period. After that, I will be at home everyday except for exams (I have 5 or 6 of them). The exam period ends at the end of June. During this period, I should be able to work or average approximately 4 hours a day.

In July and Augustus, I have nothing else than the GSoC. I don't go to holidays this year, and I don't mind working during the week-ends (programming is my passion and working on KDE is very fun!)

About me

I'm a Belgian student in Computer Science. This year, I end my third and last year of Bachelor and I hope to graduate at the bachelor level.

My website has a fairly extensive summary of all the projects (mostly open source) on which I worked for the past 6 years. There is also a page giving more details about what I like to do.

Regarding KDE, I have many years of experience with Qt and C++ development. I use KDE as my primary desktop environment for 5 years and I always followed closely its development (I have Planet KDE constantly open and I'm subscribed to kde-core-devel, kde-devel, kdevelop, kde-multimedia and nepomuk-devel).

Last year, my GSoC project was about implementing a new query parser for Nepomuk. It was a success (I think), and the parser has been ported to Baloo. I hope to have it merged in Baloo or available as a separate library relatively soon.

Junior job link: New file extractors for Nepomuk, the Nepomuk query parser, a patch to Dolphin for it to use a new Nepomuk widget, file extractors for Office 2003 files, and finally a small patch to KDevelop's QML support.

- About
- Contact
- Google+
- Blogger
- Email
- IRC
- Privacy Policy