

Фёдор Геннадьевич Сигаев

Электронное интервью.

Балдин Е.М.

конец лета/начало осени 2006

Возраст 33 года. Дипломированный физик. Профессиональный программист. Место работы: «свободный художник». Домашняя страничка: <http://www.sigaev.ru/>.

Фёдор Сигаев постоянно улучшает качество и скорость поиска PostgreSQL. Он сделал значительный вклад в развитие GiST (Generalized Search Tree) и полнотекстовый поиск. Ему нравится развивать различные алгоритмы индексации данных.

Программированием Фёдор вплотную занялся ещё будучи студентом. Дипломная работа была посвящена построению математической модели межпланетного магнитного поля и модель оказалась аналитически невычислимой, поэтому существенный объём диплома занимала программная часть. Окончание обучения совпало с «интересными временами», что привело к смене профессии, благо необходимые навыки уже были, а интерес к программированию наличествовал.

Евгений М. Балдин **Евгений:** Меня интересуют причины по которым вы включились в проект PostgreSQL. Как это случилось?

Фёдор Г. Сигаев **Фёдор:** Работой над PostgreSQL занялся во время работы в Рамблере (<http://www.rambler.ru>). Нужен был онлайн-полнотекстовый поиск (сейчас это модуль `tsearch2`) для новостей хранящихся в PostgreSQL. Так и пошло: сначала был создан модуль полнотекстового поиска, потом его индексная поддержка, а затем начали появляться улучшения уже в самом PostgreSQL. По мере того как разбирался с исходниками, появлялись идеи как его улучшить, какие типы данных для решения портал-ных задач необходимо добавить. Таким образом появились иерархический тип данных (`ltree`), поиск похожих слов (`pg_trgm`) и т. д.

Евгений: То, что вы заняты в таком проекте, вам не мешает заниматься основной работой?

Фёдор: В настоящее время это и является моей *основной* работой ☺. Ранее работодателям всегда удавалось объяснить/доказать необходимость моей работы над

PostgreSQL. А сейчас заказчики находят нас специально именно для доработки или усовершенствования PostgreSQL.

Евгений: Собственно говоря, чем это интересно? Что такого притягательного в индексах?

Фёдор: Всегда любил ковыряться поближе к основам. В программировании мне интересны алгоритмы, а не, условно говоря, пользовательские интерфейсы. А потроха базы данных — это алгоритмы и идеи.

В основу PostgreSQL заложена очень мощная идея — table-driven database. То есть, все операторы, типы, индексы и т. д. не вшиты в код, а доступны через системные таблицы. Например, для каждого типа есть функции преобразования из человеко-читаемого вида в машинный и обратно. Названия этих функций просто лежат в системной таблице, хранящей описания типов данных. Таким образом, PostgreSQL намного более легко расширяем дополнительными модулями без перекомпиляции самой базы данных. Быстрый результат приятен вдвойне.