

## Пресс-релиз компании MeraLabs

(17.05.2008)

Исследователи из компании MeraLabs, дочерней компании известной группы компаний MERA разработали и запатентовали новую технологию построения бесконечно расширяемых хранилищ электронных документов.

Генеральный директор MeraLabs и президент MERA Дмитрий Пономарев заявил, что разработанная технология в корне отличается от известных принципов построения реляционных и объектных баз данных. «Сегодня мы предлагаем новую архитектуру построения распределенных систем хранения и эффективного поиска электронных документов, представленных в форме XML файлов. Мы назвали эту архитектуру «Mera rEvolutionary Repository Architecture» («MERA»). Архитектура позволяет хранить документы на неограниченно расширяемом числе компьютеров, объединенных TCP/IP сетью, и обеспечивает релевантный быстрый поиск по многим критериям: ключевым словам, позиционным идентификаторам, объектам, нечетким запросам и т.п.»

Как прокомментировал особенности архитектуры MERA научный руководитель, технический директор MeraLabs профессор Владимир Крылов, «Мы положили в основу архитектуры исследованный нами математический объект – метризованный граф со свойствами «тесного мира». Такой подход позволил отказаться от централизованного индексирования и применить полностью распределенную систему взаимных ссылок всех документов, содержащихся в хранилище. Помещая в систему хранения новый документ, мы просто связываем его системой ссылок на уже имеющиеся документы, используя специальные алгоритмы. Каждый документ тем самым получает собственный справочник ссылок.

При появлении запроса на поиск документов с некоторыми заданными параметрами, формируется XML документ, имеющий аналогичную внутреннюю структуру и получивший название «поисковая маска». Эта поисковая маска помещается в хранилище по

тем же самым принципам что и все остальные документы, и иницируется поиск хранимых XML документов, наиболее «близких» к документу-маске. Для этого анализируются локальные ссылки, ассоциированные с каждым из «друзей» маски и осуществляется переход к следующему документу в хранилище. Далее анализируются следующие «друзья» и осуществляется следующий переход. Поиск завершается, когда достигается либо полное совпадение с маской, либо выбирается целое множество документов с заданной степенью «близости» к запросной «маске». Все они могут считаться релевантными данному запросу. Чтобы размер каждого локального справочника (число «друзей» документа) всегда оставалось малым даже для гигантских размеров хранилища, и длина пути поиска всегда была небольшой, и потребовалось привлечь экзотическую математическую теорию».

В компании MeraLabs считают, что разработанная архитектура позволит решить многие задачи построения огромных и динамичных баз данных, связанных с хранением описаний различных сущностей реального мира. Например, любые системы инвентаризации объектов недвижимости, товарных запасов, оборудования могут быть реализованы надежным образом и неограниченно расширяться.

«Мы способны предложить систему хранения, оперативного обновления и поиска описаний, например, всех автотранспортных средств или оружия, связать ее с устройствами считывания радиочастотных меток и тем самым во многом снизить потери от ненадлежащего учета таких объектов» — заметил Дмитрий Пономарев.

«Реестровые службы различных направлений могут быть построены на основе нашей архитектуры. Системы сопровождения жизненного цикла изделий, библиотеки мультимедийного контента, служба имен в глобальной сети Интернет – все это потенциальные потребители разработанной нами технологии».

Нижний Новгород, май 2008