

Руслан Богатырев

DB2 Stinger: на пути к автономному компьютерингу

Леон Кацнельсон (IBM, директор по конкурентным технологиям при департаменте информационного управления) отвечает на вопросы “Мира ПК”

В октябре 2001 г. корпорация IBM провозгласила новый манифест ИТ-индустрии — концепцию автономного компьютеринга [1]. В ее основе лежит идея создания адаптивной ИТ-инфраструктуры [2], способной автоматически приспосабливаться к изменяющимся условиям. Параллели с автономной нервной системой и живыми организмами восходят к проекту IBM по созданию цифровой иммунной системы для киберпространства (Digital Immune System for Cyberspace, [3]) и движению IBM eLisa. При этом все они связаны общими авторами. Гораздо менее известен факт проведения Джоном фон Нейманом подобной аналогии с миром нейронов в первом публичном источнике, задавшем архитектуру современных компьютеров — “First Draft on a Report on the EDVAC” (30 июня 1945 г.). Спустя шесть десятилетий идеи воплощения адаптивности компьютерных систем по образу и подобию живых организмов находят благодатную почву во флагманских продуктах IBM. С провозглашением концепции автономного компьютеринга “Голубой гигант” делает важный шаг на пути перевода разработок в области искусственного интеллекта в плоскость реального бизнеса.

Автономный компьютеринг — подход к созданию самоуправляемых компьютерных систем (self-managed computing systems), предусматривающих минимальное воздействие со стороны человека. Термин восходит к автономной нервной системе человека, которая управляет ключевыми функциями организма без участия сознания. (Из терминологического словаря IBM Research.)

Итак, каким образом идеи автономного компьютеринга нашли свое воплощение в новом продукте IBM DB2 UDB 8.2 (проект Stinger) и что уже сейчас предлагают ИТ-индустрии исследовательские и производственные подразделения IBM? На эти вопросы отвечает **Леон Кацнельсон** (IBM Toronto Laboratory).



Начиная с 2003 г. специалисты все активнее обсуждают тему IBM Stinger (DB2 Universal Database) против Microsoft Yukon (SQL Server 2005), говоря об упреждающем ударе IBM по созданию полноценной поддержки баз данных для платформы .NET. Какова в этом отношении специфика Stinger? Сколь глубока интеграция Stinger с IBM Rational XDE Developer, Microsoft Visual Studio и Eclipse?

Нас не совсем правильно понимают, когда говорят, что DB2 (<http://www-306.ibm.com/software/data/db2/udb/edition-ese.html>) и Yukon (<http://msdn.microsoft.com/msdnmag/issues/04/02/YukonBasics/default.aspx>) конкурируют на платформе .NET. СУБД DB2 имеет большое конкурентное преимущество: у нас единственная в мире СУБД, которая может делать разработку для .NET, а сама при этом способна работать на любой платформе, включая Linux, мэйнфреймы, Windows. Другая подобная многоплатформная СУБД — это Oracle. Но средства разработки в Oracle для .NET — это почти ничего. Что касается Yukon, то в определенной мере мы с ним конкурируем. Наша реализация для .NET достаточно глубокая. Простой пример: один из важных аспектов Yukon — хранимые CLR-процедуры (CLR stored procedures). Когда 2-3 года назад корпорация Microsoft объявила об этом, для разработчиков это считалось одним из самых главных моментов. Мы с этим согласны. Понятно, что когда разработчик занимается программированием на Visual Basic, он не хочет переходить на T-SQL. У нас, например, есть хранимые процедуры на Java, но разработчик, использующий, скажем, C#, далеко не всегда хочет переходить ради этого на Java. Поэтому мы уже сейчас даем такую возможность (которую Yukon планирует предоставлять в середине 2005 г.) — работать с хранимыми процедурами в виртуальной машине CLR, поддерживающей множество различных языков программирования. Кроме того, мы сделали полную интеграцию в Visual Studio .NET, чтобы прямо там можно было разрабатывать подобные хранимые процедуры. В этом смысле мы конкурируем где-то с SQL Server, включая Yukon. Мы считаем, что у нас уникальная позиция. Мы здесь конкурируем не только с Oracle, у которого такой интеграции просто нет, но даже с Microsoft, которая не может с IBM сравниться в отношении количества и разнообразия поддерживаемых нами платформ.

Мы считаем, что разработчики не обязаны иметь степень доктора философии в DB2. Вместо этого мы пытаемся подогнать СУБД и ресурсы, которые она предоставляет, так, чтобы они выглядели абсолютно естественно для разработчика в его среде. Если этот разработчик использует, например, Visual Studio .NET, то мы нацеливаемся на то, чтобы все операции, которые он должен выполнять в своем ежедневном рабочем цикле, выполнялись из Visual Studio .NET.

Появились ли в Stinger новые средства для поддержки Web-сервисов?

Несмотря на то что в области БД все мы молимся на язык SQL, для разработчиков это “необходимое зло”, они работают с ним, потому что вынуждены работать, но у них душа не лежит к этому. Мы это понимаем. Они смотрят на другие технологии. Например, Java-программисты хотят работать с JavaBeans-компонентами. Их не интересуют хранимые процедуры в SQL. Кроме того, сейчас все более популярной становится SOA-архитектура. У ИТ-директоров она находится на первом месте. И потому в этом мире многие программисты рассматривают архитектуру как создание ПО через сборку определенных модулей из Web-сервисов. В DB2 мы даем возможность пользоваться этим подходом вместо SQL. Звучит непривычно. Но это реальность.

Как заявляется, в Stinger реализована поддержка автономного компьютеринга (autonomic computing). Сама эта концепция была выдвинута IBM несколько лет назад. Весной 2004 г. в Нью-Йорке прошла первая International Conference on Autonomic Computing, где наряду с IBM приняли участие также Microsoft и HP. Означает ли это, что идея адаптивного, саморегулируемого ПО становится все более важной среди лидеров рынка?

Мы стали заниматься данным направлением гораздо раньше, чем официально это объявили. У нас было движение eLiza и такой проект, как SMART (Self-Managing And Resource Tuning — Р. Б.). Идея простая. Мы не собираемся заменить всех роботами. Просто квалифицированные кадры в области БД довольно дефицитные и дорогостоящие. Поэтому если мы можем помочь нашим администраторам БД тратить свое время на более нужные задачи вместо рутины, это надо сделать. В области автономного компьютеринга (<http://www.ibm.com/research/autonomic>) мы выделяем четыре аспекта: Self-Configuring (самонастройка), Self-Healing (самокоррекция), Self-Optimizing (самооптимизация) и Self-Protecting (самозащита). В Stinger мы сделали определенный упор на это, и нам нравится, что конкуренты подхватили нашу идею (Microsoft Dynamic Systems Initiative и HP Adaptive Enterprise — Р. Б.).

**IBM и автономный компьютеринг.
События 2004 г.**

- 30 ноября** — IBM и Fujitsu Limited объявили о сотрудничестве в сфере разработки открытых стандартов для автономного компьютеринга.
- 29 октября** — IBM выпустила Autonomic Computing Toolkit 2.0, новую версию инструментария для автономного компьютеринга.
- 20 октября** — IBM объявила об успехах внедрения автономного компьютеринга на примере трех компаний — Whirlpool, Service Delivery for Distribution Company и The Credit Union of Texas
- 30 сентября** — Выпущен новый продукт IBM DB2 Information Integrator, обеспечивающий доступ к критически важной бизнес-информации и опирающийся на средства поддержки автономного компьютеринга.
- 9 сентября** — IBM объявила о выпуске DB2 Universal Database 8.2 (проект Stinger). Среди новых средств, обеспечивающих поддержку автономного компьютеринга, — IBM Learning Optimizer и IBM Design Advisor.
- 3 августа** — IBM передала Apache Software Foundation ориентированную на Java реляционную СУБД Cloudscape объемом около 500 тыс. строк исходных текстов. В рамках открытых исходных текстов она выступает под именем Derby.
- 24 июня** — IBM объявила о выпуске WebSphere Extended Deployment, отвечающего за обеспечение автоматической оптимизации производительности корпоративного программного обеспечения и аппаратных средств.
- 16 февраля** — IBM выпустила Autonomic Computing Toolkit, ориентированный на работу в инструментальной среде Eclipse.
- 12 января** — IBM установила новый рекорд в США по числу патентов, полученных одной компанией в течение года. В 2003 г. их количество составило 3415. Одной из наиболее значимых инноваций эксперты признали патент на компьютерную систему с самокоррекцией (self-healing computer system).

Средства самонастройки баз данных, введенные в Stinger, пока слабо известны рядовым пользователям баз данных. О какой самонастройке идет речь? Насколько в связи с этим изменяются задачи администрирования БД?

Например, мы можем делать полную автоматизацию ежедневной работы БД... Все, что мы просим от администратора, — указать время, окно, когда он считает, что именно данный диапазон — самое оптимальное время для работ по поддержке БД. DB2 пытается вписать соответствующие процессы в это окно. Многие думают, что это просто планировщик работ, расписание. Но это не расписание. Например, DB2 решает, нужно ли вообще это делать. Другой пример — реорганизация БД. DB2 Stinger сам решает, нужно ли ее проводить, основываясь на заданных критериях.

Здесь проходит очень зыбкая граница между пакетным выполнением по расписанию и выполнением определенных операций без участия администратора БД, когда система сама принимает решение. Наверное нужны какие-то подсказки от администратора, чтобы он выставлял цели, приоритеты и т. д. ?

Отвечу одним предложением. Наша стратегия — это управление на основе регламентов (policy-based management). Стратегия наших конкурентов — графики и расписание. Регламент означает, что надо задать только параметры, а система сама подберет выход под эти параметры. Именно это мы развиваем. То, что вы видите сейчас в Stinger, — лишь первый шаг в данном направлении.

Прагматический аспект понятен. То, что средства автономного компьютеринга, встроенные в Stinger, высвобождают руки администратора БД, хорошо. А как это достигается? В какой мере идея автономного компьютеринга захватывает область искусственного интеллекта?

Наша задача — построение инфраструктур, а не самих приложений. Мы занимаемся инфраструктурами, которые сами себя адаптируют под приложения. Мы не занимаемся искусственным интеллектом в том смысле, что переписываются сами программы. Хотя, например,

наш оптимизатор переписывает SQL-запросы в более эффективном виде, которые быстрее выполняются.

Другими словами, вы воспринимаете программу, которую написал SQL-программист, как некую спецификацию, и по ней строите оптимальную программу?

Да, именно так. И это лишь один из примеров. Вопрос в том, как сделать так, чтобы система могла сама реагировать на изменение условий. За этим стоят огромные исследования, стоят десятки и сотни человеколет, затраченных на изучение этих проблем, решения с очень сложными формулами. Но есть и проблемы, которые математически еще не решены. Я люблю надеяться на будущее, но привык все же готовиться к настоящему. Поскольку я работаю с конкурентными технологиями, то для меня это следующий квартал, следующий год. Нам это важно, потому что выйти на рынок надо сейчас. Но при этом у нас есть целая армия исследователей, которые думают, мечтают на 10 лет вперед. Мы же все-таки завязаны на бизнес, и наша задача – это упрощение управления очень сложными системами. Чем сложнее система, тем выше потребность в ее упрощении.

Очень важный тезис. Давайте попробуем его развить. Не так давно я познакомился с работой Пола Хорна (главный вице-президент IBM Research — Р. Б.) по автономному компьютерингу [4]. В ней высказывается следующая мысль: сложность современных ИТ-систем — это серьезный вызов (grand challenge). Как это ни парадоксально, чтобы избежать сложности, утверждает Хорн, надо делать еще более сложные системы. Т. е. сложность преодолевается через сложность. Хорн подчеркивает, что следуя принципам автономного компьютеринга, надо системы еще более усложнять, но внедрять при этом элементы адаптивности. Какова Ваша точка зрения?

Сложные системы порождают сложные решения для их упрощения. Он не имел в виду, что клин клином вышибают. Т.е. чем сложнее система, тем больше необходимость ее упрощать. При этом сам процесс упрощения еще больше усложняет эти системы. Я не знаю, согласен ли будет в этом отношении Пол Хорн с моей точкой зрения: он над этим думает намного больше, чем я.

Как я понимаю, он имел в виду сложность во внутреннем исполнении, но простоту во внешнем использовании. Как только появляется адаптивная система, возникает вопрос доверия. Если администратор БД не понимает, что она автоматически делает, то доверие будет минимальным. Если он начнет это понимать (т.е. она будет ему сообщать, советовать и т.п.), тогда у него возникнет ощущение, что это помощник, которому можно доверить данный участок работы.

С этим я полностью согласен. Мы ничего не прячем и раскладываем все по полочкам: что произошло, почему произошло, какие потенциальные возможности решения этого и даем возможность администратору БД прогонять пробные сценарии (what-if), чтобы достичь его цели. Потому что, как бы мы ни стремились улучшить технологию, всегда остается человеческий фактор. Он математикой не решается. Это очень и очень важно. Какой бы ни была технология, какая бы ни стояла математика за этой технологией, всегда человек должен находиться под контролем или быть уверен, что может довериться системе. А степень доверия определяется его опытом, сколько раз он прошел через это.

Давайте вернемся к началу интервью, но на несколько на ином уровне. Для тех, кто работает с DB2, с Informix, вопрос перехода на Stinger не стоит. Это новая версия и естественный путь развития. Для тех же, кто работал с SQL Server 2000, с Visual Studio .NET, сейчас возник серьезный зазор. Далеко ушли вперед привычные им средства разработки, но реальной, достойной их СУБД на данный момент нет. Теперь она появляется в лице DB2 (как вариант). При этом существует расхожее мнение, что DB2 ориентирована преимущественно на мэйнфреймы, что ценовая точка входа в технологию довольно значительная — несколько десятков тысяч долларов. Как быть?

С точки зрения сложившегося стереотипа, вы абсолютно правы. IBM в 2004 г. сделала большой подарок сообществу открытых исходных текстов (open source community). В Apache Foundation мы передали Java-ориентированную СУБД Derby (Cloudscape; дистрибутив см. на Мир ПК-диск — Р. Б.). Она абсолютно бесплатная. При этом с ее помощью помимо Java можно разрабатывать приложения БД для .NET. Cloudscape — это бренд IBM. Мы не берем за нее деньги, только за поддержку. А Derby — проект-инкубатор в Apache DB, лежащий полностью в рамках открытых исходных текстов (<http://incubator.apache.org/derby/>). Как на это смотрят разработчики, использующие Visual Studio .NET? С технической стороны переход с SQL Server на DB2 очень

простой. Я не хочу сказать, что автоматический — что волшебной палочкой махнул, и все получилось, но у нас есть инструментарий по миграции, который позволяет преобразовывать данные, хранимые процедуры и т. д. В то же время интерфейсы, например, ADO.NET на DB2 и SQL Server абсолютно одинаковый. Все, что нужно сделать, это глобальный поиск и замену объектов SQL на объекты DB2.

Такова техническая сторона. Но есть и психологическая, о которой вы сказали. В восприятии многих DB2 — это мэйнфрейм. Но это не так. Наша задача подобный стереотип сломать, изменить. С технической точки зрения, по нашему мнению, SQL Server 2000 просто нельзя даже сравнивать с DB2 Stinger. SQL Server 2000 дал все, что мог. Yukon задержался. Вы знаете, что он год назад уже должен был появиться, но все никак не появится. И пока Yukon не появится, разработчики действительно застряли где-то на полпути, на технологии пятилетней давности. В Yukon будут интересные вещи, но для многих людей это тупик, потому что, как ни крути, это Windows. Он привязан к этой платформе. Это его сила, но это и огромный его недостаток. Поэтому перед нами не стоит техническая задача. Наша задача, особенно в таких странах, как Россия, — поломать стереотип восприятия того, что люди слышали о DB2 с эпохи 1980-х годов.

Литература

1. Jeffrey Kephart, David Chess "The Vision of Autonomic Computing" // IEEE Computer, January 2003.
http://www.research.ibm.com/autonomic/research/papers/AC_Vision_Computer_Jan_2003.pdf
2. Леонид Черняк "Адаптируемость и адаптивность" // Открытые системы, 2004, № 9.
<http://www.osp.ru/os/2004/09/030.htm>
3. Руслан Богатырев "Совершенно секретно, или Всемирная электронная нервная система" // Мир ПК, 1998, № 4.
<http://www.osp.ru/pcworld/1998/04/128.htm>
4. Paul Horn "Autonomic Computing: IBM's Perspective on the State of Information Technology" — IBM Research, October 2001.
http://www-3.ibm.com/autonomic/pdfs/autonomic_computing.pdf