

ПРОБЛЕМЫ ИНТЕГРАЦИИ БАЗ ДАННЫХ В ГЕТЕРОГЕННЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Андропов А. В. (Университет ИТМО)

Научный руководитель – к.т.н., доцент Дергачев А.М.
(Университет ИТМО)

Аннотация В рамках данного исследования рассматриваются проблемы, с которыми можно столкнуться при решении задач интеграции информационных систем. При этом подразумевается, что сами системы различны по своему составу (гетерогенны) и используют для хранения информации базы данных.

Введение. Современные информационные системы таких информационных гигантов как Google, Яндекс, Amazon насчитывают сотни или даже тысячи сервисов, разработанных для разных программных платформ и выполняющихся под управлением различных операционных систем. При этом сервисы не могут существовать обособленно от других сервисов системы, им необходимо взаимодействовать друг с другом. Для организации взаимодействия на программном уровне могут использоваться удаленные вызовы процедур, системы обмена сообщениями и другие средства. Зачастую в интеграции информационных систем участвуют и базы данных. Также стоит отметить, что с развитием интернет-ориентированной деятельности предприятий задача интеграции систем распространилась на среду Интернет. Это привело к расширению организации взаимодействия между сервисами одной системы к взаимодействию между целыми информационными системами. Все эти факторы существенно влияют на задачу интеграции и порождает некоторое множество проблем.

Основная часть. Большинство проблем интеграции связаны с гетерогенностью информационных систем. Так различные подсистемы могут быть реализованы на разных языках программирования и использоваться различные СУБД.

Большинство СУБД имеют драйвера, позволяющие приложениям взаимодействовать с базами на различных платформах. Стоит отметить, что взаимодействие всегда проходит через некоторого посредника будь то драйвер или другое программное средство. На сегодняшний день широко распространены технологии объектно-реляционного отображения (ORM), позволяющие абстрагироваться от деталей взаимодействия с конкретными СУБД. Такая технология существенно облегчает интеграцию, но порождает проблему в необходимости поддерживать этот уровень абстракции постоянно в актуальном состоянии.

Решать возникающие проблемы интеграции можно следующими способами:

- 1) стандартизация метаданных о структуре СУБД;
- 2) интеграция посредством конверторов;
- 3) интеграция с помощью сервисов;
- 4) интеграция с использованием метамоделей;
- 5) интеграция на уровне бизнес-процессов;

Не во всех СУБД есть возможность получить информацию о схеме базы данных в реальном времени, такую как названия таблиц, типы данных и т.д. Такая возможность значительно упростила бы разработку, особенно при использовании технологии ORM. Создать такую возможность в рамках одной СУБД недостаточно, необходимо вывести это на уровень стандарта.

В качестве следующего способа — создание дополнительных подсистем/конверторов, которые будут обращаться в обе СУБД, возможно разными способами.

Создание сервисов с единым интерфейсом для каждой отдельно выбранной СУБД. Таким образом система обращается к сервисам, а не к самой СУБД.

Под метамоделью понимается информационная модель более высокого уровня абстракции чем модель предметной области. Она описывает не отдельную задачу, а широкий круг задач, задаёт правила обработки данных и управления бизнес-процессами. При наличии такой модели каждая система определяется набором параметров этой модели. Соответственно интеграция между системами представляет собой перевод данных из одной модели в другой.

Интеграция на уровне бизнес-процессов более подходящая для предприятий, так как их деятельность состоит по большей части из бизнес-процессов, а не приложений, баз данных и платформ. Идея, лежащая в основе такой интеграции, заключается в создании сценария некоторого бизнес-процесса и описания в нём операций взаимодействия пользователей с системами и систем между собой

Каждое решение имеет свои преимущества и недостатки, поэтому выделить одно общее решение довольно трудно, но при выборе стоит учитывать тот факт, что в нынешнее время большинство информационных систем являются распределёнными.

Выводы. Таким образом, основные проблемы интеграции в том, что сами информационные системы не рассчитаны на взаимодействие с другими системами, а также работу в гетерогенной среде. Эти проблемы могут проявляться как в выборе языка программирования, так и при выборе СУБД. Решение этих проблемы в большинстве информационных систем уникально и не следует какому-либо общераспространённому подходу, что может влечь за собой повторение ошибок уже существующих решений. Стоит заметить, что наличие единого стандартизированного реестра для информационных систем помогло справиться с этой проблемой и сделало бы интеграции одних систем в другие более гибкими и динамичными.

Андропов А.В. (автор)

Подпись

Дергачев А.М. (научный руководитель)

Подпись