

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.383

Бренько
Сергей Юрьевич

Метод поддержки распределенных транзакций в микросервисной
архитектуре

АВТОРЕФЕРАТ

на соискание академической степени
магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

подпись магистранта

Научный руководитель
Смолякова О. Г.
к.т.н., доцент

подпись научного руководителя

Минск 2019

КРАТКОЕ ВВЕДЕНИЕ

Практически каждый запрос, обрабатываемый современным корпоративным приложением, выполняется в рамках транзакции базы данных. Разработчики корпоративных приложений используют фреймворки и библиотеки, которые упрощают управление транзакциями. Большинство из этих фреймворков предоставляют декларативные механизмы управления транзакциями, в результате чего написание транзакционной бизнес-логики в монолитной архитектуре становится относительно простой задачей.

На первый взгляд традиционные распределенные транзакции кажутся очень привлекательными. С точки зрения разработчика, они имеют практически ту же модель программирования, что и локальные транзакции. Однако из-за того, что координатор транзакции является единой точкой отказа, чрезмерно высокого объема коммуникации между координатором и участниками транзакции, наличия блокировок, отсутствия поддержки большинством современных технологий и слишком большого влияния на доступность данных, использование двухфазного или трехфазного коммита кажется не самым подходящим решением для приложений, построенных на микросервисной архитектуре.

Практически каждое современное приложение в той или иной степени хранит данные. Согласованность этих данных подразумевает, что все экземпляры приложения постоянно представлены с одним и тем же набором значений данных. Этот подход иногда называют строгой согласованностью данных.

В наши дни облачные приложения обычно используют данные, которые распределены по нескольким хранилищам данных. Управление и поддержание согласованности данных в такой среде может стать критическим аспектом системы, особенно если принять во внимание проблемы параллелизма и доступности данных, которые могут возникать. Зачастую это приводит к тому, что необходимо пожертвовать строгой согласованностью данных в пользу их доступности. Это означает, что при проектировании приложения стоит уделить внимание идее согласованности в конечном итоге и признать, что данные, используемые приложением, могут не всегда находиться в согласованном состоянии.

Чтобы решить проблему поддержания согласованности данных в микросервисной архитектуре, приложение должно использовать механизм, основанный на концепции слабо связанных, асинхронно-коммуницирующих сервисов.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка метода поддержки распределенных транзакций в микросервисной архитектуре, обладающего высокими характеристиками производительности, способного обеспечить преемственность новых технологий и предоставляющего возможность достижения согласованности данных.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Проанализировать традиционные подходы к поддержке распределенных транзакций и оценить их применимость для микросервисной архитектуры;
2. Выявить критерии оценки качества распределенных транзакций;
3. Разработать метод поддержки распределенных транзакций, применимый для микросервисной архитектуры и соответствующего необходимым критериям;
4. Предоставить решения для потенциальных проблем разработанного метода;
5. Оценить применимость разработанного метода к практической разработке и соответствие необходимым критериям.

Объектом исследования является микросервисная архитектура построения программных приложений.

Предметом исследования являются методы поддержки распределенных транзакций в микросервисной архитектуре, которые обеспечивают высокий уровень согласованности данных.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность реализации метода поддержки распределенных транзакций, который обеспечит максимально эффективное использование аппаратных средств распределенной системы в совокупности высоким уровнем согласованности данных и производительности. Строгая согласованность данных для большинства случаев избыточна, в связи с чем согласованность в конечном итоге выглядит предпочтительнее и может оказать положительное влияние на характеристики системы.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры «Программное обеспечение информационных технологий» по теме «Разработка моделей, методов, алгоритмов, повышающих показатели проектирования, внедрения и эксплуатации программных средств для перспективных платформ обработки информации, решения интеллектуальных задач, работы с большими массивами данных и внедрение в современные обучающие комплексы» (ГБ № 16-2004, № ГР 20163588, научный руководитель НИР – Н. В. Лапицкая).

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя О. Г. Смоляковой, заключается в формулировке целей и задач исследования.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на Республиканской научно-технической конференции «Информационные технологии и системы: проблемы, методы, решения» (Минск, Беларусь, 2018).

Опубликованность результатов диссертации

По теме диссертации опубликовано 2 печатных работ, из них 1 статья в рецензируемом издании, 1 работа в сборниках трудов и материалов конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников, списка публикаций автора и приложений. В первой главе представлен анализ классических решений для управления распределенными транзакциями в микросервисной архитектуре и выявлены критерии оценки качества распределенных транзакций. Вторая глава посвящена разработке метода поддержки распределенных транзакций в микросервисной архитектуре, который обеспечивает необходимый уровень согласованности данных. В третьей главе проведена оценка применимости разработанного метода к практической

разработке, а также оценка соответствия разработанного и существующих методов выявленным критериям качества.

Общий объем работы составляет 66 страниц, из которых основного текста – 49 страниц, 24 рисунков на 13 страницах, 2 таблицы на 1 странице и списка использованных источников из 36 наименований на 3 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** описана потребность в наличии распределенных транзакций при разработке приложений, основанных на микросервисном подходе, обрисованы основные подходы к организации распределенных транзакций, проведен анализ их применимости для микросервисной архитектуры и выявлены критерии для оценки качества распределенных транзакций.

Среди традиционных подходов к распределенным транзакциям можно выделить протоколы двухфазного и трехфазного коммитов.

Протокол двухфазного коммита заключается в наличии фаз транзакции, таких как запроса на фиксацию и фиксации, а также координатора, который указывает участникам транзакции какую фазу стоит выполнять.

Протокол трехфазного коммита является своего рода эволюцией протокола двухфазного коммита, а основной мотивацией к его созданию является уменьшение количества блокировок. Данный протокол состоит из трех фаз: проверка возможности проведения операции, выполнение предварительного коммита, фиксация изменений. Уменьшение количества блокировок происходит за счет упразднения блокировок, возникающих при отказе одного из участников транзакции или координатора.

Данные протоколы трудно применимы к микросервисной архитектуре, так как обладают рядом недостатков, которые имеют высокую значимость для данного архитектурного подхода. К таким недостаткам относятся: отказоустойчивость, вызванная тем, что координатор является единой точкой отказа; наличие блокировок; повышение коммуникации между компонентами по сети; преобладание современных технологий становится невозможной, так как они не поддерживают данные протоколы; снижение доступности системы, так как такие распределенные транзакции представляют собой форму синхронного межпроцессорного взаимодействия.

Выделенные критерии для оценки качества распределенных транзакций, следующие: необходимость в масштабировании, отказоустойчивость, надежность, пропускная способность, производительность, объем вносимых задержек, преемственность технологий и влияние на доступность.

Вторая глава посвящена методу поддержки распределенных транзакций, который отлично применим к микросервисной архитектуре. При этом описаны применяемый подход к согласованности данных, сущность метода, и потенциальные проблемы. Для возможных проблем предложены решения.

Как подход к согласованности данных предложена концепция согласованности в конечном итоге, основанная на том, что данные в системе могут находиться в несогласованном состоянии во время выполнения связанных операций по их обновлению.

В качестве метода поддержки распределенных транзакций и поддержания согласованности данных в микросервисной архитектуре, предлагается использовать механизм, основанный на концепции слабо связанных, асинхронно-коммуницирующих сервисов – Фазы. При данном подходе каждая распределенная транзакция разбивается на последовательность небольших локальных транзакций.

В качестве механизма координирования шагов фазы предложены варианты хореографического управления последовательности шагов (участники фазы подписываются на события системы и обновляют данные соответствующим образом; координатор отсутствует) и управление на основе оркестровки (координатор фазы определяет следующего участника и указывает ему какую операцию выполнять).

Отсутствие свойства изолированности является еще одной проблемой при применении фаз. При данной проблеме возможны следующие аномалии данных: потерянные обновления, грязные чтения и неповторяющиеся чтения. Для решения проблемы предложен ряд контрмер: семантические блокировки, коммутативные обновления, повторные чтения и прочие.

В качестве методов коммуникации предложены различные принципы асинхронного взаимодействия.

Также предоставлены решения для транзакционного обмена асинхронными сообщениями. К этим решениям относятся опрос таблицы базы данных и применение отслеживания журнала событий.

В третьей главе рассмотрена практическая реализация системы, основанной на использовании фаз. Для основных проблем, возникающих при разработке описаны соответствующие решения. Система построена по модульно-функциональному принципу, когда основные функциональные

действия реализуются отдельными классами со строго обозначенной зоной ответственности.

Согласно результатам сравнительной оценки соответствия необходимым критериям, подход, основанный на использовании фаз, является более предпочтительным для микросервисной архитектуры, чем применение протоколов двухфазного и трехфазного коммита. Как результат можно утверждать, что использование фаз влечет за собой положительное влияние на эффективность использования приложения и повышает его доступность, производительность, отказоустойчивость и сокращает необходимость в его масштабировании.

ЗАКЛЮЧЕНИЕ

1. Проведена оценка необходимости применения распределенных транзакций при разработке корпоративных программных приложений, основанных на микросервисной архитектуре. При этом проведено сравнение традиционного подхода, основанного на применении монолитной архитектуры, и подхода с использованием микросервисов при практической реализации транзакционной логики.

2. Исследованы традиционные подходы и методики, применяемые при реализации распределенных транзакций, такие как протоколы двухфазного и трехфазного коммита. Также было выполнено их сравнение, описаны основные достоинства и недостатки, предоставлены графические нотации, отображающие принципы работы данных алгоритмов.

3. Проведен анализ применимости протоколов двухфазного и трехфазного коммита для задач согласованности данных при использовании микросервисной архитектуры. Выделены основные проблемы, которыми обладают данные протоколы. К таким проблемам относятся:

- Низкие показатели отказоустойчивости из-за того, что координатор транзакции является единой точкой отказа;
- Наличие блокировок при выполнении транзакции;
- Повышенный объем коммуникации между участниками транзакции;
- Низкая преемственность современных технологий, таких как базы данных NoSQL и системы обмена асинхронными сообщениями, при разработке из-за того, что эти технологии не поддерживают протоколы двухфазного и трехфазного коммита;
- Снижение показателей доступности из-за фактического использования синхронного межпроцессного взаимодействия.

4. Выявлены критерии оценки качества распределенных транзакций. Описана необходимость в соответствии этим критериям и их влияние на эффективность программного продукта. К таковым критериям относятся: необходимость в масштабировании, отказоустойчивость, надежность, пропускная способность и производительность, вносимые задержки, преемственность технологий, влияние на доступность всей системы.

5. Выделена проблема согласованности данных распределенного приложения. Проведено сравнение достижимости согласованности данных при использовании монолитной и микросервисной архитектур. Описана разница между строгой согласованностью данных и их согласованностью в конечном итоге.

6. Описан механизм поддержания согласованности данных в микросервисной архитектуре, основанный на использовании на концепции слабо связанных, асинхронно-коммуницирующих сервисов. Данным механизмом являются Фазы, которые представляют собой последовательность локальных транзакций, посредством которых происходит достижение согласованности данных в конечном итоге.

7. Обрисована необходимость использования компенсирующих транзакций как механизма восстановления системы и достижения согласованности данных по результатам каждой фазы. Описаны основные проблемы, которые могут возникнуть при реализации компенсирующих транзакций и предложены их решения.

8. Выделены подходы к координации фаз. К таковым подходам относятся хореографическое управление фазами и управление на основе оркестровки. Описаны их преимущества и недостатки. Даны рекомендации к практическому применению обоих подходов.

9. Выявлена проблема отсутствия свойства изолированности транзакций при использовании фаз. Приведен список возможных аномалий данных, таких как потерянные обновления, грязные чтения и неповторяющиеся чтения. Также описан ряд контрмер, которые помогают в избегании возможных аномалий и компенсации отсутствия изолированности. Среди данных контрмер можно выделить семантическую блокировку, коммутативные обновления, повторное чтение значений, использование файлов версии и оценку по значению.

10. Охарактеризованы асинхронные методы коммуникации между участниками транзакции, применяемые при реализации фаз. К таковым методам относятся коммуникации по принципу “запрос – ответ” или “запрос – асинхронный ответ”, коммуникация по принципу “публикация – подписка”

и “публикация – асинхронные ответы”, а также использование принципа односторонних уведомлений.

11. Описана проблема транзакционного обмена асинхронными сообщениями. Выделены два решения для задач публикации событий: путем опроса таблицы базы данных и с применением отслеживания журнала транзакций. Описаны основные достоинства и недостатки данных подходов. Даны практические советы по реализации.

12. Выполнена оценка применимости фаз к практической разработке. При этом был разработан опытный образец приложения, использующий фазы, описаны классы реализации и их зоны ответственности, показаны каналы коммуникаций и взаимодействия. Также были показаны практические реализации решений основных проблем, возникающих при использовании фаз, таких как: необходимость в компенсирующих транзакциях и координирование фазы (хореографический подход и подход с использованием оркестратора).

13. Выполнена оценка соответствия необходимым критериям традиционных подходов, таких как протоколы двухфазового и трехфазового коммита, и подхода, основанного на использовании фаз. При этом было проведено сравнение по каждому из необходимых критериев. Как результат можно утверждать, что использование фаз влечет за собой положительное влияние на эффективность использования приложения и повышает его доступность, производительность, отказоустойчивость и сокращает необходимость в масштабировании.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Бренько С. Ю. Управление распределенными транзакциями в микросервисной архитектуре / Бренько С. Ю. // Республиканская научно-техническая конференция «Информационные технологии и системы: проблемы, методы, решения» – Минск, 2018.

2. Бренько С. Ю. Применение фаз при разработке распределенных приложений / Бренько С. Ю. // Журнал «Молодой ученый» №23 (261) – июнь 2019 – с. 6 – 9.