

НЕБЛОКИРУЮЩИЙ ПРОТОКОЛ ДВУХФАЗНОЙ ФИКСАЦИИ ТРАНЗАКЦИЙ

А. А. Скачихин

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: alexei.skachykhin@live.com

В работе предлагается модификация протокола двухфазной фиксации транзакций устойчивая к отказам координатора. Отказоустойчивость достигается за счёт добавления резервного координатора. Улучшенный протокол применим в распределённых системах, где время исполнения транзакций является критичным и внедрение более сложного протокола не удовлетворяет этому требованию.

ВВЕДЕНИЕ

Задачи связанные с обработкой данных и проведением вычислений, всё чаще требуют мощностей целого кластера компьютеров для их успешного выполнения [1]. В распределённых системах остаётся актуальной задача реализации ACID транзакций. За последние десятилетия, исследователями было предложено множество протоколов обеспечивающих достижение консенсуса, необходимого для реализации транзакционности [2]. Одним из них является протокол двухфазной фиксации транзакций, далее ПДФТ. Основное преимущество ПДФТ заключается в малом количестве сетевых сообщений необходимых для достижения консенсуса. В свою очередь его существенный недостаток в том, что при отказе узла координатора, исполнение транзакции блокируется на продолжительный период времени или вовсе не может быть закончено. Предложенная в данной работе модификация, позволяет устранить этот недостаток.

I. ПРОТОКОЛ ДВУХФАЗНОЙ ФИКСАЦИИ ТРАНЗАКЦИЙ

Основная идея ПДФТ заключается в том, что достижение консенсуса происходит за два шага (рис. 1). На первом шаге координатор рассылает сообщения менеджерам ресурсов (Request-to-prepare). Менеджеры ресурсов инициируют локальные транзакции (например сохранение в реляционную базу данных). Когда транзакции были выполнены, но ещё не зафиксированы, менеджер ресурсов отправляет подтверждение (Vote) координатору. Второй шаг начинается, когда координатор получает ответы от всех менеджеров ресурсов, он инициирует фиксацию (Commit) или отмену (Abort) транзакции путём отправки сообщения менеджерам ресурсов. Менеджеры ресурсов фиксируют/отменяют локальную транзакцию и высылают подтверждение координатору (Acknowledgement). Всего для достижения консенсуса требуется $3n$ сообщений, где n - количество менеджеров ресурсов.

Основные сложности в процессе достижения консенсуса связаны с отказами узлов во время исполнения распределённой транзакции. В ПДФТ координатор является единой точкой отказа системы, ведь именно координатор принимает решение о том, считать транзакцию успешной или нет. Если отказ координатора происходит после первой фазы, то менеджеры ресурсов продолжают удерживать блокировку на локальных ресурсах до получения запроса на фиксацию от координатора. Если в системе не предусмотрен механизм обеспечения отказоустойчивости к отказам координатора, то для снятия блокировки может потребоваться вмешательство человека, что является неприемлемым для многих задач решаемых в распределённых системах.

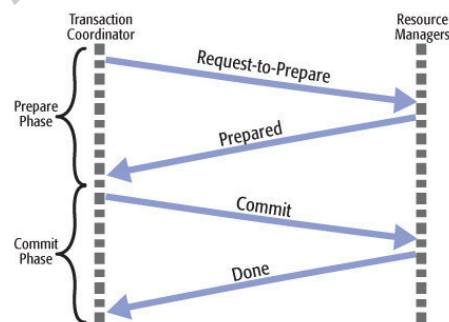


Рис. 1 – ПДФТ

В качестве решения этой проблемы выступает протокол трехфазной фиксации транзакций, где координатор может быть выбран путём голосования и в случае отказа координатора распределённые транзакции не блокируются [2]. Подобное поведение достигается путём отправки дополнительных $2n$ сообщений, где n - количество менеджеров ресурсов. Такое решение сказывается на времени исполнения распределённой транзакции.

По мнению автора в ряде задач ПДФТ является более предпочтительным в связи с меньшим количеством сообщений необходимых для достижения консенсуса, а проблему отказов координатора видится возможным решить с помощью добавления резервного координатора следя-

шего за состоянием основного координатора. В случае отказа основного координатора, резервный занимает его место.

II. ПРОТОКОЛ ДВУХФАЗНОЙ ФИКСАЦИИ С РЕЗЕРВНЫМ КООРДИНАТОРОМ

В основе предлагаемого усовершенствования ПДФТ лежит дополнительный узел, который в случае отказа основного координатора выполняет его роль (рис. 2).

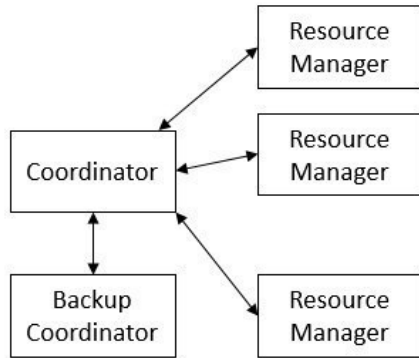


Рис. 2 – ПДФТ

Для того, чтобы резервный координатор мог продолжить транзакцию, а основной координатор в случае восстановления не пытался повторить транзакцию, предусмотрен механизм выбора лидера. Таким образом только один координатор может является основным [3]. Основной и запасной координаторы имеют доступ к общему журналу транзакций и таким образом запасной координатор узнаёт на какой фазе находится исполнение транзакции при отказе основного координатора. Между собой координаторы общаются посылкой периодических ring сообщений. В случае если ответ на ring сообщение не приходит за ограниченный период времени, происходит замена координатора. В качестве механизма замены координатора предлагается использовать DNS, таблицу маршрутизации или очередь сообщений без явного указания адресата.

Так как отказ координатора может произойти на любом этапе исполнения распределённой транзакции, важно определить поведение резервного координатора для всех состояний системы после отказа.

- Отказ до начала первой фазы - наиболее простой случай, при котором по истечении допустимого времени ответа на ring сообщение запасной координатор повторно иницирует первую фазу;
- Отказ во время первой фазы - уже полученные голоса (votes) запасной координатор может считать из журнала транзакций, если какие-то голоса не были обработаны координатором до его отказа, запасной координатор высасает запрос в менеджеры ресурсов для повторного получения голо-

сов. Тут стоит отметить, что менеджеры ресурсов должны хранить голоса некоторое время на случай отказа координатора;

- Отказ во время первой фазы после получения всех голосов - запасной координатор иницирует вторую фазу;
- Отказ во время второй фазы - поведение аналогично отказу во время первой фазы;
- Отказ после второй фазы - в этом случае запасной координатор опрашивает ресурсные менеджеры о статусе их локальной транзакции и на основании ответов принимает решение считать ли распределённую транзакцию успешной.

Добавление запасного координатора позволило существенно уменьшить вероятность блокировки транзакции:

$$P = P_1^2,$$

где P_1 - вероятность отказа координатора. Время T в течение которого транзакция остаётся заблокированной равняется времени допустимого ожидания ответа на ring сообщение и может конфигурироваться исходя из нужд решаемой задачи.

Выводы

В работе было предложено улучшение протокола двухфазной фиксации транзакций, которое позволяет частично решить проблему блокировки распределённых транзакций в случае отказа координатора. Метод основан на резервировании и поэтому существенно уменьшает вероятность блокировки. Улучшенный протокол применим в распределённых системах, где время исполнения транзакций является критичным и внедрение более сложного протокола не удовлетворяет этому требованию. Предложенное улучшение не устраняет факт блокировки, так как возможна ситуация отказа одновременно двух координаторов. Дальнейшим развитием идеи может стать отказ от отправки ring сообщений. Это становится возможным если менеджеры ресурсов будут обладать знанием о наличии резервного координатора.

СПИСОК ЛИТЕРАТУРЫ

1. Armbrust, M. A View of Cloud Computing, Communications of the ACM / M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica and M. Zaharia // Communications of the ACM. - 2010
2. Poonam, S. An Extended Three Phase Commit Protocol for concurrency control in Distributed Systems / S. Poonam, P. Yadav, A. Shukla // International Journal for computer applications. - 2011. - Vol. 21, № 10. - P. 213-215
3. Helmy, T. Dynamic Load-Balancing Based on a Coordinator and Backup Automatic Election in Distributed Systems / T. Helmy // International Journal of Computing Information Sciences. - 2011. - Vol. 9, № 1. - P. 24-28