

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

УДК 004.75

Романович
Евгений Александрович

Алгоритм достижения консенсуса в распределенных системах

АВТОРЕФЕРАТ

на соискание степени магистра технических наук

по специальности 1–40 80 02 «Системный анализ, управление и обработка информации»

Научный руководитель

Матвеев В.В.

Минск 2020

ВВЕДЕНИЕ

Системы и приложения центров обработки данных работают в высокодинамичных средах. Они расширяются за счет использования ресурсов дополнительных серверов и растут, и сокращаются в зависимости от нагрузки. Сбои в работе серверов и сетей также являются обычным явлением: около 2–4% дисководов выходят из строя каждый год, примерно так же часто происходит сбой серверов, и десятки сетевых соединений ежедневно выходят из строя в современных центрах обработки данных.

В результате системы должны иметь дело с серверами, которые во время нормальной работы могут останавливаться и восстанавливаться. Они должны реагировать на изменения и автоматически адаптироваться в течение нескольких секунд; сбои, которые заметны для людей, как правило, не приемлемы. Это серьезная проблема в современных системах; Обработка сбоев, координация, обнаружение услуг и управление конфигурацией - все это сложно в таких динамических средах.

К счастью, распределенный консенсус может помочь с этими проблемами. Консенсус позволяет совокупности машин работать как единая группа, которая может пережить неудачи некоторых ее членов. В консенсусной группе сбои обрабатываются принципиальным и проверенным способом. Поскольку консенсусные группы являются высокодоступными и надежными, другие компоненты системы могут использовать консенсусную группу в качестве основы для собственной отказоустойчивости. Таким образом, консенсус играет ключевую роль в создании надежных крупномасштабных программных систем.

Алгоритм Raft один из самых популярных алгоритмов нахождения консенсуса в распределенных системах на данный момент. Он намного проще в понимании относительно своих конкурентов, хоть и не проигрывает в эффективности, однако еще недостаточно изучен и может быть улучшен.

На основании всего вышеизложенного, в магистерской диссертации поставлена основная задача на изучение, повышение эффективности и стабильности работы алгоритма Raft.

Для достижения поставленной цели необходимо выполнить следующее:

- 1) Изучить алгоритм достижения распределенного консенсуса Raft в распределенных системах, а также прочие доступные алгоритмы.
- 2) Выяснить, какие недостатки и ограничения он имеет и предложить соответствующее возможное решение представленных проблем.
- 3) Реализовать и проверить достоверность и эффективность предлагаемых модификаций.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

На данный момент существует множество подходов к разработке различного рода систем и комплексов. Однако в последнее время все большую популярность приобретает понятия распределенности системы как на физическом уровне так и на программном. Распределенность позволяет взаимодействовать с аппаратной, а в некоторых случаях и с программной частью на расстоянии, что невероятно удобно. Так же в понятие распределенная система входит клиент-серверная архитектура использующаяся сейчас повсеместно. В свою очередь распределенные системы вносят свою сложность и определяют дополнительные ограничения. Одной из таких проблем является проблема достижения согласования или нахождения консенсуса.

Степень разработанности проблемы

Алгоритмов нахождения консенсуса в распределенных системах не так и много. Абсолютным лидером на данный момент является алгоритм Raft разработанный Диего Онгаро и Джоном Оустерхаутаром в 2014 году. Это алгоритм пользуется большой популярностью среди разработчиков при решении проблемы согласованности для распределенных облачных кластеров, баз данных или надежных систем. В свою очередь алгоритм был представлен исключительно с академической точки зрения и не был достаточно протестирован в реальных условиях. Усовершенствования, представленные в данной работе, являются обязательными при реализации и применении данного алгоритма на практике.

Цель и задачи исследования

Целью магистерской диссертации является улучшение алгоритма консенсуса Raft для распределенных асинхронных систем.

Объектом исследования являются компьютерные распределенные асинхронные системы.

Предметом исследования является алгоритм достижения распределенного консенсуса Raft.

Для достижения поставленной цели необходимо было решить следующие задачи:

1. Провести обзор и анализ алгоритма raft.
2. Провести анализ требований, предъявляемых к системе, и принципа работы алгоритмов распределенного консенсуса;

3. Разработать предлагаемый алгоритм и провести соответствующие эксперименты.

Теоретическая и методологическая основа исследования

При проведении исследования и написании диссертации использованы научные публикации, техническая документация и интернет-источники, посвященные вопросам компьютерного проектирования электронных систем, разработке многофункциональных клиент-серверных приложений, принципам работы и использования алгоритмов достижения консенсуса.

Для решения поставленных задач использованы следующие методы исследования: анализ, синтез, обобщение, сравнение, логический и графический методы.

Научная новизна

Научная новизна заключается в создании модификаций алгоритма распределенного консенсуса Raft.

Теоретическая значимость работы заключается в анализе метода достижения согласования систем в распределенной асинхронной среде.

Практическая значимость работы заключается в разработанном алгоритме, усовершенствующем работу кластеров системы.

Апробация диссертации и информация об использовании ее результатов

Результаты исследований, вошедших в диссертацию, докладывались на 56-ой научной конференции аспирантов, магистрантов и студентов БГУИР (г. Минск, Республика Беларусь, 2019–2020 г.).

СОДЕРЖАНИЕ РАБОТЫ

Первая глава диссертации показывает в полной мере что такое задача нахождения распределенного консенсуса и ограничения накладываемые на систему. Рассматривается CAP теорема в рамках распределенных систем, а также алгоритмы согласования, предшествующие алгоритму Raft: 2-х и 3-х фазные коммиты, алгоритм Пакос и алгоритм репликации с отметкой.

Вторая глава диссертации сфокусирована на самом алгоритме Raft. Описан подход к реализации, а также сам протокол. Четко определены методы и свойства разных состояний в системе: последователя, кандидата и лидера. Представлено объяснение того, как происходят выборы лидера и как работает согласование журнала в системе. В заключительном разделе главы описаны дополнительные условия, обеспечивающие безопасность работы алгоритма и системы в целом. Также на протяжении всей главы приведены примеры нарушения корректности работы протокола при игнорировании условий и ограничений, налагаемых на систему. В завершении главы приводятся гарантии для распределенного консенсуса, описанные протоколом.

В основе третьей главы лежит основная работа, проделанная в данной диссертации – представлены возможные усовершенствования алгоритма. Для облегчения работы системных администраторов, а также более эффективной работы системы предложено ввести универсальный уникальный идентификатор в систему, позволяющий вводить в систему новые узлы, основываясь на уже существующих узлах. В свою очередь это также облегчает задачу восстановления кластеров при неблагоприятных условиях работы сети.

Следующей оптимизацией является дополнение протокола алгоритмом преждевременного голосования, который выполняется перед переходом в статус кандидата. Суть состоит в том, что сервер может переключиться на кандидата только в том случае, если преждевременное голосование выполнено успешно, в противном случае он должен дождаться следующего таймаута выборов. Заключительная модификация протокола – это устранение уязвимости алгоритма, приводящее к гонке за лидерство. Решение, в свою очередь, возможно благодаря алгоритму преждевременного голосования: последователи должны отклонять новых лидеров, если с их точки зрения существующий лидер по-прежнему функционирует правильно, что означает, что они получили сообщение от лидера раньше, чем истекло время ожидания новых выборов.

В заключительной главе описана реализация алгоритма с модификациями, а также приведены результаты экспериментов. Проведенные эксперименты доказали достоверность и эффективность описанных улучшений.

ЗАКЛЮЧЕНИЕ

Основной целью данной магистерской диссертации является рассмотрение и разработка усовершенствований алгоритма распределенного консенсуса Raft.

Raft хорошо подходит для применения в практических системах: он описан достаточно подробно для реализации без дальнейшей доработки, он решает все основные проблемы в системе и является эффективным. Raft использует архитектуру, которая более применима для построения систем: консенсус часто определяется как согласие на одно значение, но в Raft его определили в терминах реплицированного журнала, поскольку это необходимо для построения реплицированного конечного автомата. Raft эффективно управляет реплицированным журналом, используя своего лидера; Для совершения запроса требуется только один раунд RPC от лидера.

Кроме того, эта работа предлагает усовершенствования, которые делают протокол еще более эффективным. Введение унифицированного уникального идентификатора позволяет администраторам видеть, при нарушении работы кластера, какие журналы требуют репликации с помощью последнего слепка данных, а какие могут быть восстановлены самим протокол более эффективно. Также введение фазы досрочного голосования решает проблему частного случая, когда сервера попадают в своеобразную гонку за лидерство. В свою очередь это повышает доступность и долгосрочность работы сервера, так как лидер будет оставаться лидером настолько это возможно, и ненужные выборы будут игнорироваться.

Во время обучения в магистратуре принял участие в 56-й научной конференции аспирантов, магистрантов и студентов БГУИР.

На основании всего изложенного, считаю, что поставленная цель достигнута в полной мере.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

[1-А] Романович, Е.А. Способы обеспечения эргономичности и эффективности автоматизированного рабочего места сотрудника банка / Мышковец А.В. // Материалы 56-ой научной конференции студентов, магистрантов, аспирантов УО «Белорусский государственный университет информатики и радиоэлектроники» Минск, БГУИР, 2020.

Библиотека БГУИР