

Научная статья

УДК 004.75

DOI 10.25205/1818-7900-2022-20-2-18-26

О количественных метриках согласованности и доступности в распределённых базах данных

Антон Игоревич Злыгостев¹, Борис Николаевич Пищик²

^{1,2}Новосибирский государственный университет

²Федеральный исследовательский центр информационных и вычислительных технологий

Новосибирск, Россия

¹a.zlygostev@g.nsu.ru

²pishchik.boris@ya.ru

Аннотация

Рассматривается вопрос одновременной достижимости согласованности и доступности. Сравниваются подходы к определению максимально возможных уровней доступности для различных моделей согласованности. Формулируется необходимость введения количественных метрик для этих двух показателей, рассматриваются работы с количественными оценками согласованности. Подтверждается вывод об актуальности данной проблемы, и предлагаются направления для моделирования параметров прикладной задачи при помощи численных экспериментов.

Ключевые слова

распределённые базы данных, теорема CAP, согласованность, доступность

Для цитирования

Злыгостев А. И., Пищик Б. Н. О количественных метриках согласованности и доступности в распределённых базах данных // Вестник НГУ. Серия: Информационные технологии. 2022. Т. 20, № 2. С. 18–26. DOI 10.25205/1818-7900-2022-20-2-18-26

On Quantitative Metrics for the Consistency and Availability in Distributed Databases

Anton I. Zlygostev¹, Boris N. Pischik^{1,2}

Novosibirsk State University

Novosibirsk, Russian Federation

2 Federal Research Center for Informational and Computational Technologies

Novosibirsk, Russian Federation

¹a.zlygostev@g.nsu.ru

²pishchik.boris@ya.ru

Abstract

This article considers simultaneous attainability of consistency and availability. The approaches to determining the maximum possible levels of accessibility for different consistency models are compared. The necessity of introducing quantitative metrics for these two indicators and review the prior articles with quantitative estimates of consistency is advocated. The importance of this problem is confirmed, and directions for modeling the parameters of the applied problem using numerical experiments are proposed.

Keywords

distributed databases, CAP Theorem, Consistency, Availability

For citation

Zlygostev A. I., Pischik B. N. On Quantitative Metrics for the Consistency and Availability in Distributed Databases. Vestnik NSU. Series: In-formation Technologies, 2022, vol. 20, no. 2, p. 18–26. (in Russ.) DOI 10.25205/1818-7900-2022-20-2-18-26

© Злыгостев А. И., Пищик Б. Н., 2022

Введение

Распределённой считается вычислительная система, состоящая из нескольких вычислительных узлов, каждый из которых находится в некотором *состоянии*; и *связей* между этими узлами, которые позволяют узлам обмениваться *сообщениями*.

Популярны несколько причин для перехода от монолитных систем к распределённым:

- повышение надёжности, или *резервирование* – аналогично механическим и иным инженерным системам, можно пытаться собрать надёжную систему из ненадёжных компонентов, заменяющих друг друга в случае сбоя;
- повышение производительности, или *распараллеливание* – распределяя вычислительную нагрузку на несколько узлов, можно повышать общую производительность (быть может, ценой дополнительных затрат на объединение результатов);
- улучшение времени отклика, или географическое *распределение* – передвигая вычислительные узлы «ближе» к клиентам в терминах количества сетевых переходов, можно уменьшить наблюдаемые сетевые задержки.

Все эти факторы становятся все более значимыми в настоящее время, когда вычислительные услуги получают все более широкое распространение: одновременно растет и нагрузка (количество пользователей), и географический охват аудитории, и само разнообразие сервисов. Однако во всех этих случаях наличие состояния у вычислительных узлов затрудняет реализацию корректной обработки пользовательских запросов. В идеальной сети коммуникация между узлами происходит без потерь, поэтому можно поддерживать у пользователей иллюзию «единого» состояния, скрывая факт распределённости.

В реальной сети отказам подвержены как узлы, так и связи между ними. Это означает, что распространение изменений состояния может задерживаться, приводя к потерям согласованности. В таких случаях при обработке пользовательского запроса узел системы может столкнуться с дилеммой – либо отказаться выполнять запрос, либо выполнить его с риском выдать не вполне корректный результат.

С точки зрения параметров распределённой системы, отказ от выполнения запроса означает снижение *доступности*, или способности системы выполнять работу, ради которой она была спроектирована; а продолжение обработки, невзирая на корректность, означает отказ от *согласованности*.

Согласованность в данном контексте означает *выполнение всех ограничений целостности*, сформулированных для прикладной задачи. Более узкая трактовка термина «согласованность» – отсутствие противоречий в состоянии данных между узлами распределённой системы – малоинтересна для пользователя, т. к. его интересует только поведение системы, наблюдаемое «извне». При приобретении билетов в кинотеатр совершенно неважно, насколько одинаковы данные о проданных местах в различных узлах кластера, если в итоге обеспечивается гарантия «не более одного билета на один сеанс на одно место».

Для разработчиков коммерческих распределённых систем или приложений вопросы выбора между согласованностью и доступностью играют очень важную роль, т. к. имеют финансовые последствия. Большинство коммерческих сервисов предоставляют так называемые *гарантии уровня сервиса*, SLA, которые ведут к выплате штрафов (или снижению выручки), если простой приложения превышает заранее оговоренные ограничения. Рассогласование данных также может приводить к убыткам – например, стандартной практикой решения вопросов с избыточной продажей авиабилетов является предоставление повышенного класса обслуживания и другие вознаграждения «пострадавшим» пассажирам.

Предсказание подобных финансовых последствий требует количественных оценок для показателей доступности и согласованности. Например, при проектировании системы резервирования авиабилетов нас будут интересовать ответы на вопросы типа «какая доля запросов

будет встречать отказ из-за сбоев компонентов системы» и «какую долю пассажиров придется везти бизнес-классом из-за несогласованности узлов».

При проектировании системы банковского обслуживания нас будут интересовать аналогичные вопросы о доле отказов, а также о матожидании сумм/сроков технической задолженности при снижении остатков на счетах ниже оговоренных лимитов.

Целью предстоящей работы является исследование *количественных* соотношений между параметрами согласованности и доступности.

Обзор подходов

Вопросы доступности и согласованности привлекали внимание авторов как в области систем управления базами данных, так и в области системного программирования с момента их зарождения.

В 1976 году Эсваран и другие вводят понятие *транзакция* как группа операций, которая переводит базу данных из одного согласованного состояния в другое [1].

Напрямую из этого понятия следует термин *сериализуемость* – такой порядок параллельного выполнения индивидуальных операций в транзакциях, при котором результат эквивалентен некоторому последовательному исполнению этих же транзакций. Простым и мощным следствием из сериализуемости исполнения является гарантия сохранения согласованности, доказываемая по индукции. На следующие двадцать лет понятие сериализуемости становится золотым стандартом согласованности в мире СУБД. Пападимитриу публикует подробный разбор теории сериализуемости [2] для одномашинных реализаций.

Ротни и Гудман поднимают вопросы работоспособности распределённой обработки данных в случае сетевых сбоев еще в 1977 году [3], т.е. до формализации реляционной алгебры.

Тем временем, Бернштейн с Гудманом публикуют в 1980-1983 серию работ, посвященных проблемам достижения сериализуемости в реплицируемых СУБД, в которых сосредотачиваются на сбоях узлов сети [4; 5; 6]. Византийские сбои узлов и сбои связей (т.е. разделение сети) в них не рассматриваются. В этих работах формулируется понятие *сериализуемости единственной копии* (one-copy serializability), которое является наиболее сильным требованием к согласованности в распределённых базах данных.

В 1985 году Дэвидсон, Гарсия-Молина и Скин публикуют работу [7], посвященную классификации стратегий по обработке разделений сети, и обсуждению методик выбора подходящей стратегии. В ней авторы отмечают, что цели доступности (минимизации нарушений работы системы в случае сбоев) и корректности не являются независимыми, и требуют некоторых компромиссов. «Так как достижение обеих целей явно невозможно, одна из них или обе должны быть в некоторой степени ослаблены, в зависимости от требований приложения. Снижение доступности достаточно прямолинейно: можно просто запрещать определённые транзакции в определённых местах [сети]. В свою очередь, снижение корректности обычно требует глубокого знания того, что представляет собой информация в базе данных, как приложения манипулируют этой информацией, и чего будет стоить отмена/компенсация/исправление несогласованности» [7, стр. 343].

Классификация подходов предлагается по двум критериям:

- предпочтения между доступностью и согласованностью: *пессимистический* подход фокусируется на согласованности и предполагает, что если в системе есть возможность совершить операции, несогласуемые между разделёнными репликами, то она обязательно будет использована; *оптимистический* подход предпочитает доступность, и предполагает, что за время разделения сети несогласуемые изменения данных, скорее всего, не произойдут;
- способ определения корректности: *синтаксический* подход основан на сериализуемости единственной копии, и проверяет сериализуемость исключительно путём сравнения наборов

данных, читаемых и изменяемых каждой транзакцией; семантический подход рассматривает семантику данных или транзакций при определении корректности.

Интересно отметить, что авторы называют «оптимистическим» протокол, предложенный Дэвидсон в её докторской диссертации 1982 года [8], несмотря на то, что он все же добивается сериализуемости путем отмены транзакций после восстановления связи между сегментами сети. Таким образом, формально он должен находиться среди пессимистических подходов в этой классификации.

Более того, работа [8] рассчитывает долю отмененных транзакций; это могло бы дать количественную оценку *доступности* системы, построенной на основе предлагаемого протокола, если бы была возможность отличить отмены, связанные с нарушением прикладных ограничений, от отмен, вызванных распределённостью системы. Интуитивно понятно, что клиент, имея на счету одну тысячу рублей, может запросить снятие наличными десяти тысяч, и получение отказа в данном случае не является сбоем системы – это корректное поведение системы, заданное прикладными требованиями. А вот если клиент провел серию транзакций, которая сама по себе не нарушает никаких инвариантов, и была бы успешно выполнена одно-машинной реализацией, но впоследствии часть транзакций пришлось отменить из-за сбоев в коммуникациях между узлами банковской системы, то это можно считать сбоем.

В работе Дэвидсон также рассмотрено несколько других протоколов, как оптимистических, так и пессимистических; к сожалению, для них не упоминаются какие-либо количественные оценки итоговой доступности или согласованности – и вообще, авторы явно не задаются подобной целью, несмотря на многократные упоминания «повышения доступности» или «частичного отказа от согласованности».

Подводя итоги, авторы отмечают, что «важно рассматривать эффективность предложенных стратегий, хотя получение обоснованных оценок для компромиссов эффективности затруднено. В одних случаях это следует из сложностей построения адекватной модели; в других – из того, что механизм сильно зависит от приложения».

В 1996 году Бирман и Фридман публикуют небольшую статью [9], где формально доказывают невозможность достижения сериализуемости для некоммутативных операций, которые выполняются во время разделения сети; они предлагают комбинированный подход, в котором коммутативные операции могут исполняться без ограничений, а некоммутативные могут исполняться при условии готовности отказаться от их результатов после восстановления связи между сегментами, что эквивалентно подходу Дэвидсон, упомянутому ранее.

В последующие годы коммутативные операции привлекают значительное внимание, благодаря своей способности к конвергенции операций, выполняемых параллельно, независимо от их порядка – например, [10].

В 1999 году Фокс и Брюер публикуют короткую статью [11], которая в значительной мере сформировала ландшафт постреляционных систем обработки данных на следующие двадцать лет. В статье формулируется «строгий CAP-принцип», декларирующий невозможность построения распределённых систем, которые одновременно достигают *сериализуемости единственной копии* и *высокой доступности*, при условии возможности сетевых сбоев между частями такой системы. В статье также упоминается и мягкий CAP-принцип: «Эти примеры предполагают, что существует также Мягкий Принцип CAP, который нам еще предстоит точно охарактеризовать: чем более сильные гарантии даны для любой пары из строгой согласованности, высокой доступности, или устойчивости к разделению, тем слабее гарантии можно дать относительно третьей». Строгий вариант принципа в дальнейшем фигурирует в литературе под названием «гипотеза Брюера», а мягкий вариант более нигде не упоминается. Основными понятиями статьи являются сбор (*harvest*) и выход (*yield*), которые определяют количественные характеристики доступности (*выход* равен доле успешно обработанных запросов и в точности соответствует инженерному показателю доступности, традиционно измеряемому в количестве

цифр 9 после запятой), и согласованности (*сбор* характеризует полноту данных, полученных в ответе на запрос).

В 2002 году Гилберт и Линч публикуют статью [12], в которой дают утверждениям гипотезы Брюера строгие формулировки и доказывают их, после чего принцип получает широкую известность под названием «теорема CAP».

Эта теорема завладевает умами большинства разработчиков на многие годы, несмотря на очевидные недостатки, отмеченные авторами прямо в работе: чрезмерная строгость трактовки понятий «согласованность» и «доступность».

Стоит отметить, что в понятии «согласованности», использованном в этой работе, нет никаких упоминаний ограничений целостности. Фокус внимания перемещается с соблюдения ограничений, заданных предметной областью, на способность системы поддерживать непротиворечивость состояний различных реплик.

По-видимому, подразумевается автоматическое выполнение гарантий целостности, сформулированных еще в 1976 году [1], в случае возможности гарантировать сериализуемость выполняемых операций.

В последующие годы можно отметить быстрый рост количества работ, посвященных различным аспектам частичной согласованности. Понятие «транзакции» уходит на второй план, и большинство авторов, вслед за Гилбертом и Линч, сосредотачивают свою трактовку согласованности на изменениях атомарного регистра. Даже для такого упрощенного описания распределенной системы предлагаются модели согласованности с различными уровнями строгости, как правило, охарактеризованные качественным образом.

Одним из немногих исключений из этого правила является работа 2009 года [13], в которой Фекити и др. дают количественную оценку влияния снижения уровня изоляции в РСУБД на частоту нарушений согласованности.

Важной работой в этом направлении является, опубликованная в 2011 году Голабом, Ли и Ша [14], – в ней исследуются количественные оценки согласованности – «отсталость» реплики от актуальной версии в терминах времени и количества версий, а также «частота нарушений», определяющая долю нарушений согласованности в истории изменения регистра.

В 2012 году Бейлис и другие развивают результаты работы [14], вводя *вероятностно ограниченную отсталость* как количественную оценку несогласованности данных [15].

Тем временем Махаджан, Алвизи и Далин [16] уточняют результаты Гилберта и Линч, доказывая, что наиболее строгой моделью согласованности, достижимой в постоянно доступной распределенной системе с возможностью разделения, является *причинная согласованность реального времени*.

В 2013 году Бейлис и другие возвращаются к вопросу о согласованности данных при обработке транзакций, включающих несколько операций над различными элементами данных [17]. Они рассматривают богатый набор различных моделей транзакционной согласованности, и находят точную верхнюю границу транзакционной согласованности, достижимую при тех же формулировках понятия доступности и разделения сети, что и у Линч с Гилбертом. В частности, из четырех уровней изоляции ANSI доказуемо достижимы два, включая наиболее популярный в промышленных базах данных Read Committed. Ценным вкладом данной работы является применение полученных результатов к тестовой задаче TPC-C – золотому стандарту обработки транзакций, который с девяностых годов является основной метрикой производительности для реляционных СУБД. Большинство прикладных задач OLTP, традиционно проблемных для распределенной реализации (продажа авиабилетов, банковские переводы, биржевые операции), сводимы к тем же примитивам, что и задача TPC-C. Три из пяти видов транзакций в TPC-C совместимы с высокой доступностью и разделением сети, четвертую можно сделать совместимой, слегка модифицировав ее условия, и только реализация пятой представляет собой существенную сложность.

Это означает, что можно построить распределённую систему, выполняющую ТРС-С, и в случае разделения она сможет продолжать принимать транзакции четырех из пяти типов без потери согласованности, и только один вид транзакций придется отложить до восстановления связи.

В 2015 году Клеппманн публикует статью [18] с критикой CAP-теоремы. Основной объект критики – определения, выбранные Гилбертом и Линч, аргумент – их непригодность в практических сценариях. «Требуется лучшая основа для описания компромиссов. Такая основа должна быть проста в понимании, соответствовать интуитивным ожиданиям большинства людей, и использовать определения, которые формальны и верны», пишет Клеппманн. Он предлагает заменить понятие «разделение сети» на «задержку связи», полагая, что разделение рано или поздно закончится (а если разделение произошло навсегда, то некоторые операции будет невозможно завершить). В дальнейшем он вводит в этой статье понятие «чувствительность к задержкам», которое характеризует поведение системы в случае таких задержек. Это – значительный шаг вперед по сравнению с наивной моделью Линч и Гилберта, оперирующей неопределёнными временами разделения.

Перрин в своей книге 2017 года [19] подробно исследует различные модели согласованности. К сожалению, количественным характеристикам согласованности и доступности там внимания не уделено; вместо этого автор сосредотачивается на технических средствах обеспечения выбранного уровня согласованности в распределённых приложениях. Нельзя не отметить широту охвата решений, известных на момент публикации, и наглядность иллюстраций. По аналогичному, хотя и более простому пути, идут Майтер, Шоллерс и Де Мютер [20] в своей библиотеке CAPtain, которая предлагает конструировать распределённые системы из объектов двух классов – «доступных» и «согласованных».

Гонзалес де Мендевил и другие в статьях, опубликованных в 2015 [21] и 2019 [22], уточняют и расширяют список моделей согласованности, совместимых с доступностью при разделении сети.

Ли, Батени и другие в препринте, опубликованном в сентябре 2021 [23], предлагают количественные метрики для доступности и согласованности. Точнее, они вводят понятия количественной *несогласованности* – интервала времени, в течение которого распределённая система может проявлять нарушения согласованности, и количественной *недоступности* как интервала времени, в течение которого система реагирует на запрос пользователя. Система с нулевой несогласованностью всегда находится в согласованном состоянии, а если несогласованность ограничена сверху произвольной константой, то можно говорить о «согласованности в конечном итоге» (eventual consistency). Нулевая недоступность соответствует мгновенному отклику системы на запросы пользователя. Если недоступность невозможно ограничить сверху некоторой константой, то некоторые запросы могут вовсе не получить ответов, и система будет демонстрировать потерю доступности. Важным вкладом этой работы является доказательство CAL-теоремы, или соотношения между определенными таким образом параметрами недоступности и несогласованности:

$$\bar{A}_i = \max \left(O_i, \max_{j \in N} (\mathcal{L}_{ij} - \bar{C}_{ij}) \right), \text{ где:}$$

$\bar{A}_i$ – недоступность i -го процесса,

O_i – задержка обработки i -го процесса,

$\mathcal{L}_{ij}$ – наблюдаемое запаздывание j -го процесса относительно i -го процесса (включая задержку обработки O_i),

$\bar{C}_{ij}$ – несогласованность процессов i и j .

Недостатком предложенных метрик является трудность их сопоставления с принятыми в инженерной и деловой практике показателями. В частности, общепринятая трактовка доступности является безразмерной величиной в диапазоне от 0 до 1; предложенная авторами

недоступность имеет размерность времени, и переход от нее к традиционной метрике неочевиден.

Постановка задачи

Таким образом, тема согласованности и доступности все еще вызывает интерес исследований; задача анализа соотношения количественных метрик согласованности и доступности, упоминавшаяся у Брюера и Клеппмана, до сих пор далека от решения. Более того – всего некоторые работы в исследуемой области используют какие-либо количественные метрики для любого из обсуждаемых критериев. При этом ни в одной из работ не предлагаются метрики, соответствующие привычной инженерно-экономической интерпретации. Этот недостаток ограничивает практическую применимость подходов, предложенных до сих пор.

По-видимому, одним из самых серьезных затруднений является специфичность прикладных задач. Если для *доступности* у нас есть привычная количественная характеристика, соответствующая доле необслуженных пользовательских запросов (или доле *времени*, в течение которого система не была способна обрабатывать запросы), то для *согласованности* трудно ввести такую метрику, которая бы не зависела от особенностей прикладной задачи.

Можно попробовать выбрать некоторую предметную область, и ввести метрику согласованности, пользуясь особенностями конкретной задачи. Исследования зависимости количественных метрик для согласованности и доступности для этой задачи могут иметь прикладное значение для разработчиков, которые строят решения в той же предметной области. Кроме того, возможно, это приблизит нас к исследованию более общего случая.

В качестве прикладной задачи можно выбрать задачу о моделировании банковской системы. Банковская система состоит из набора счетов. Ее состояние полностью описывается значениями *остатков* на каждом из счетов. Остатки изменяются при помощи *транзакций* – операций перемещения денег со счета на счет. Таким образом, сумма остатков на всех счетах всегда остается неизменной.

На уровне отдельного счета вводится дополнительное требование согласованности – запрет на списание денег ниже лимита остатка счета. Это ограничение нарушает коммутативность операций снятия и пополнения счета, затрудняя реализацию банковской системы в распределённой среде.

Для оценки выполнения этого требования согласованности можно ввести метрику, которую легко интерпретировать с точки зрения бизнеса: доля времени, в течение которого остаток на счету является неотрицательным. Распределённость системы означает, что у различных ее узлов (процессов) могут быть различные представления о «текущем остатке» в каждый момент времени. Поэтому при определении согласованности мы не будем пользоваться этими представлениями, а предположим, что все транзакции можно упорядочить относительно некоторых глобальных часов, независимо от того, какой из узлов отвечает за обработку каждой из транзакций. Тогда остатком на счету в каждый момент времени мы будем называть сумму всех пополнений/списаний, которые были приняты всей системой к этому моменту:

$$B(t) = \sum_{Tr.timestamp < t} Tr.amount$$

$B(t)$ описывает некоторую функцию от t . Можно рассчитать суммарную длительность всех интервалов времени, в течение которых эта функция удовлетворяет ограничению целостности:

$$T_c = \int_T \left\{ \begin{matrix} 0, B(t) < 0 \\ 1, B(t) \geq 0 \end{matrix} \right\} dt$$

Тогда собственно целостность можно определить как отношение T_c к T – общему времени функционирования исследуемого счёта:

$$C = \frac{T_c}{T} = \frac{1}{T} \int_T \begin{cases} 0, B(t) < 0 \\ 1, B(t) \geq 0 \end{cases} dt$$

В таком определении согласованность всегда находится в диапазоне от 0 до 1, при этом для системы, не нарушающей согласованности в ни на какой промежуток времени, C будет точно равно 1.

Такая метрика будет хорошо сочетаться с традиционной инженерной трактовкой доступности, определённой в виде *доли запросов*, которые успешно обрабатываются системой.

Различные стратегии обработки разделения сети (протоколы согласования) могут приводить к различным сочетаниям значений метрик доступности и согласованности.

Рассмотрение возможных сочетаний этих значений за пределами очевидных решений «добиваемся 100 % доступности ценой потери согласованности» и «добиваемся 100 % согласованности ценой потери доступности» может представлять существенный интерес.

Список литературы / References

1. **Eswaran K. P., Gray J. N., Lorie R. A., Traiger I. L.**, The notions of consistency and predicate locks in a database system, Communications of the ACM, vol. 19, № 11, 1976.
2. **Papadimitriou C. H.**, The serializability of concurrent database updates, Journal of ACM, vol. 26, № 4, 1979.
3. **Rothnie J. B., Goodman N.**, A survey of research and development in distributed database management, Proceedings of the third international conference on Very large data bases, vol. 3, 1977.
4. **Bernstein P. A., Goodman N.**, Timestamp-based algorithms for concurrency control in distributed database systems, Proceedings of the sixth international conference on Very Large Data Bases, vol. 6, 1980.
5. **Bernstein P. A., Goodman N.**, Concurrency control in distributed database systems, ACM Computer Surveys, vol. 13, № 2, 1981.
6. **Bernstein P. A., Goodman N.**, The failure and recovery problem for replicated databases, 1983.
7. **Davidson S. B., Garcia-Molina H., Skeen D.**, Consistency in Partitioned Networks, Computing Surveys, vol. 17, № 3, 1985.
8. **Davidson S. B.**, An optimistic protocol for partitioned distributed database systems, Princeton: Princeton University, 1982.
9. **Birman K., Friedman R.**, Trading Consistency for Availability in Distributed Systems, Cornell University, Cornell, 1996.
10. **Roh H.-G., Jeon M., Kim J.-S., Lee J.**, Replicated abstract data types: Building blocks for collaborative applications, Journal of Parallel and Distributed Computing, vol. 71, 2011.
11. **Fox A., Brewer E. A.**, Harvest, Yield, and Scalable Tolerant Systems, Proceedings of the The Seventh Workshop on Hot Topics in Operating Systems, 1999.
12. **Gilbert S., Lynch N.**, Brewer's Conjecture and the Feasibility of Consistent, Available, Partition-Tolerant Web Services, ACM SIGACT News, vol. 33, № 2, pp. 51–59, June 2002.
13. **Fekete A., Goldrei S. N., Asenjo J. P.**, Quantifying Isolation Anomalies, Proceedings of the VLDB Endowment, Lyon, 2009.
14. **Golab W., Li X., Shah M. A.**, Analyzing Consistency Properties for Fun and Profit, Proceedings of the 30th annual ACM SIGACT-SIGOPS symposium on Principles of distributed computing, San Jose, 2011.
15. **Bailis P., Venkataraman S., Franklin M. J., Hellerstein J. M., Stoica I.**, Probabilistically Bounded Staleness for Practical Partial Quorums, Proceedings of the VLDB Endowment, vol. 5, № 8, 2012.
16. **Mahajan P., Alvisi L. и Dahlin M.**, Consistency, Availability, and Convergence, Department of Computer Science, The University of Texas, Austin, 2012.

17. **Bailis P., Davidson A., Fekete A., Ghodsi A., Hellerstein J. M. и Stoica I.**, Highly Available Transactions: Virtues and Limitations, Proceedings of the VLDB Endowment, vol. 7, № 3, 2014.
18. **Kleppmann M.**, A Critique of the CAP Theorem, arXiv preprint arXiv:1509.05393, 2015.
19. **Perrin M.**, Distributed Systems, Concurrency and Consistency, London: ISTE Press Ltd, 2017.
20. **Myter F., Scholliers C., De Meuter W.**, A CAPable Distributed Programming Model, Proceedings of the 2018 ACM SIGPLAN International Symposium on New Ideas, New Paradigms and Reflections on Programming and Software, Boston, 2018.
21. **Pascual-Miret L., Gonz'alez de Mend'ivil J. R., Bernab'eu-Aub'an J. M., Mu'noz-Escoi F. D.**, Widening CAP Consistency, Universitario Mixto Tecnol'ogico de Inform'atica, Universitat Polit'ecnica de Val'encia, Valencia, 2015.
22. **Mu'noz-Escoi F. D., de Juan-Mar'in R., Garc'ia-Escriv'a J.-R., Gonz'alez de Mend'ivil J., Bernab'eu-Aub'an J. M.**, CAP Theorem: Revision of its Related Consistency Models, The Computer Journal, vol. 62, № 6, pp. 943-960, 2019.
23. **Lee E. A., Bateni S., Lin S., Lohstroh M. и Menard C.**, Quantifying and Generalizing the CAP Theorem, 16 9 2021. [В Интернете]. Available: <https://arxiv.org/abs/2109.07771>. [Дата обращения: 13 10 2021].

Сведения об авторах

Злыгостев Антон Игоревич, магистрант, Новосибирский государственный университет, (Новосибирск, Россия)

Пищик Борис Николаевич, кандидат технических наук, старший научный сотрудник, Федеральный исследовательский центр информационных и вычислительных технологий, (Новосибирск, Россия); доцент кафедры компьютерных систем, Новосибирский государственный университет (Новосибирск, Россия)

SPIN code: 4297-3046

AuthorID: 13080

Information about the Authors

Anton I. Zlygostev, master's student, Novosibirsk State University (Novosibirsk, Russian Federation)

Boris N. Pishik, Candidate of Technical Sciences, senior researcher, Federal Research Center for Information and Computing Technologies (Novosibirsk, Russian Federation); Associate Professor, Department of Computer Systems, Novosibirsk State University (Novosibirsk, Russian Federation)

SPIN code: 4297-3046

AuthorID: 13080

Статья поступила в редакцию 08.04.2022;

одобрена после рецензирования 26.05.2022; принята к публикации 26.05.2022

The article was submitted 08.04.2022;

approved after reviewing 26.05.2022; accepted for publication 26.05.2022