**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 5](#_Toc105166645)

[ОБЗОР АНАЛОГОВ 8](#_Toc105166646)

[1 ОПИСАНИЕ HASH ARRAY MAPPED TRIE 11](#_Toc105166647)

[1.1 Описание структуры данных Trie 11](#_Toc105166648)

[1.2 Описание однопоточной Hash Array Mapped Trie 15](#_Toc105166649)

[1.3 Описание lock-free Hash Array Mapped Trie 23](#_Toc105166650)

[2 РАЗРАБОТКА HASH ARRAY MAPPED TRIE 28](#_Toc105166651)

[2.1 Настройка рабочего окружения 29](#_Toc105166652)

[2.2 Реализация однопоточной версии 33](#_Toc105166653)

[2.3 Реализация блокирующей многопоточной версии 38](#_Toc105166654)

[2.4 Реализация неблокирующей версии 39](#_Toc105166655)

[2.5 Решение проблемы ABA в неблокирующей версии 41](#_Toc105166656)

[2.6 Внедрение в библиотеку libcds 42](#_Toc105166657)

[3 ТЕСТИРОВАНИЕ HASH ARRAY MAPPED TRIE 43](#_Toc105166658)

[3.1 Unit-Тестирование 43](#_Toc105166659)

[3.2 Тестирование производительности 45](#_Toc105166660)

[4 ЗАКЛЮЧЕНИЕ 49](#_Toc105166661)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 51](#_Toc105166662)

[Ссылки 51](#_Toc105166663)

# 

# ВВЕДЕНИЕ

Одно из требований к современному программному обеспечению направленно на эффективное использование ресурсов вычислительных устройств. Многопоточные алгоритмы зачастую позволяют повысить производительность и в определённых случаях они оказываются быстрее однопоточных. Если в однопоточной версии единственный поток выполняет всю работу индивидуально, то в многопоточной несколько потоков делят работу на меньшие части и выполняют их одновременно (это позволяет эффективнее использовать ресурсы процессора и повышать производительность программного обеспечения).

Сложность написания таких алгоритмов заключается в том, что потоки зачастую имеют доступ к одним и тем же данным (ресурсам). Одновременная работа с данными может привести систему в несогласованное состояние. Это вызовет неправильную работу потоков и системы в целом.

Проблема некорректной работы с данными может решаться через блокировку потенциально опасных частей кода, где есть доступ к ресурсам. В таких критических местах может исполняться только один поток, что делает работу с данными безопасной. Пока такой поток работает в критической секции, другие, желающие исполниться, ожидают (останавливают свою работу). Такой подход иногда приводит к снижению производительности из-за затрат на использование критических блоки. Более того, при неправильной реализации возможна бесконечная остановка системы (взаимоблокировка), когда каждый поток входит в ожидание.

Альтернативой являются алгоритмы без блокировок (lock-free). В них потоки постоянно выполняют необходимую работу, не ожидая вход в какой-либо участок кода. Такие алгоритмы могут улучшить производительность за счёт отсутствия бесконечной остановки системы, отсутствия накладных расходов на критические секции и отсутствия ожидания потоков.

Использование lock-free подхода распространено в библиотеке libcds [1] (библиотека для многопоточных структур данных). Она включает в себя множество высокопроизводительных, неблокирующих структур на языке C++. Например, многопоточные реализации хеш-таблиц (MichaelHashMap, FeldmanHashMap) или деревьев (BronsonAvlTree). Такие структуры данных зачастую необходимы, так как они решают распространённую в программировании задачу поддержания словаря. К сожалению, во многих случаях их производительности недостаточно из-за особенностей таких структур данных. Например, хеш-таблицы в некоторые моменты снижают производительность из-за внутреннего устройства, а определённые структуры данных основанных на деревьях часто имеют большое время исполнения операций. Компромиссным решением для эффективного поддержания словаря является структура данных Hash Array Mapped Trie. Она сочетает в себе свойства и деревьев, и хеш-таблиц.

В библиотеке libcds и её аналогов Boost [2] и ConcurrencyKit [3] неблокирующая структура данных Hash Array Mapped Trie не присутствует. Поэтому появилась необходимость внедрения такой структуры данных в библиотеку libcds. Она может быть производительнее своих аналогов в задаче поддержания словаря.

Целью данной работы является повышение производительности программного обеспечения за счет реализации потокобезопасной структуры данных Hash Array Mapped Trie. Для её достижения были поставлены следующие задачи:

1. Анализ существующих структур данных
2. Реализация однопоточной версии Hash Array Mapped Trie
3. Реализация блокирующей потокобезопасной версии   
   Hash Array Mapped Trie
4. Реализация неблокирующей потокобезопасной версии   
   Hash Array Mapped Trie
5. Решение проблемы ABA в неблокирующей реализации
6. Тестирование разработанной структуры данных
7. Внедрение в библиотеку libcds

# ОБЗОР АНАЛОГОВ

При внедрении неблокирующей (lock-free) Hash Array Mapped Trie решается задача поддержания *словаря* [4], в котором каждый элемент (значение) идентифицируется уникальным ключом. Под поддержанием имеется ввиду возможность производить некоторые операции. Обычно, в этот список входят:

* Вставка значения по ключу
* Удаление значения по ключу
* Поиск значения по ключу

Важно, чтобы время исполнения каждой операции было как можно меньше, а затраты памяти на хранение словаря стремились к минимуму.

Задача поддержания такого словаря решается повсеместно: начиная от самой популярной базы данных Redis [5] и заканчивая таблицей символов в большинстве компиляторов [6]. Причём, в зависимости от специфики области применения некоторыми требованиями можно пренебречь в пользу других. Например, пожертвовать временем исполнения для сокращения затрачиваемой памяти или памятью для того, чтобы предотвратить снижение производительности. В таблице 2 показана производительность структур данных, поддерживающих такой словарь.

Таблица 2 – Структуры данных и их среднее время исполнения операций

|  |  |  |  |
| --- | --- | --- | --- |
|  | Добавление | Удаление | Поиск |
| Хеш-Таблица |  |  |  |
| Красно-Чёрное Дерево |  |  |  |
| Список |  |  |  |

Из таблицы 2 можно сделать вывод, что хеш-таблица является производительнее других структур данных по времени исполнения. Но стоит отметить [7]:

* во-первых, хеш-таблице время от времени необходимо время на изменение своей структуры (перехеширования всех элементов) и последующее расширение
* во-вторых, в хеш-таблице (с открытой адресацией) удаление ключа освобождает память, занятую при его добавлении, лишь частично.
* в-третьих, хеширование занимает время. Более того, оно происходит как минимум один раз при каждой операции с деревом.

Первое замечание ведёт к понижению производительности в случайный момент времени. Это может быть недопустимым в некоторых системах реального времени. Второе замечание показывает неэффективность   
хеш-таблиц по памяти (хранятся лишние элементы).

Красно-чёрное дерево не тратит время на перехеширование и не хранит лишних элементов при удалении, но имеет скорость выполнения операций выше, чем в хеш-таблице.

Список является самым эффективным по памяти, но скорость выполнения операций выше, в чем хеш-таблице и дереве.

Появляется необходимость в структуре данных, которая является компромиссом между быстрой хеш-таблицей и эффективным по памяти красно-чёрным деревом. Такой структурой данных является Hash Array Mapped Trie. Все операции у неё выполняются за , где – глубина дерева. В отличие от красно-чёрного дерева Hash Array Mapped Trie производит меньше дополнительных операций при выполнении вставки, удаления или поиска. Также в ней отсутствует необходимость в перехешировании. Hash Array Mapped Trie при удалении не хранит лишние элементы, как это делает хеш-таблица. И главное, Hash Array Mapped Trie может быть реализована как неблокирующая структура данных. Это даёт прирост к производительности её операции (вставка, удаление, поиск).

# 

# 1 ОПИСАНИЕ HASH ARRAY MAPPED TRIE

1.1 Описание структуры данных Trie

Идея внедряемого алгоритма была описана в статье Cache Aware Lock-Free Concurrent Hash Tries [8]. Перед тем как изложить соответствующее устройство неблокирующего Hash Array Mapped Trie, необходимо понять, как работает однопоточная версия. Это, в свою очередь, требует знаний о структуре данных Trie.

*Trie* – дерево поиска (оно же *префиксное дерево* или бор) позволяет хранить, удалять, искать ключи, а также привязанные к ним значения. Обычно, ключами выступают строки, но могут быть и числа. Чтобы добавить элемент в дерево, ключ разбивается на части. Если ключ – это строка, то разбиение зачастую идёт по символам. В случае, если ключ – это число, то на биты. Далее для каждой части создаётся репрезентирующий её узел. Например, на рисунке X показано разбиение строки «HTON»

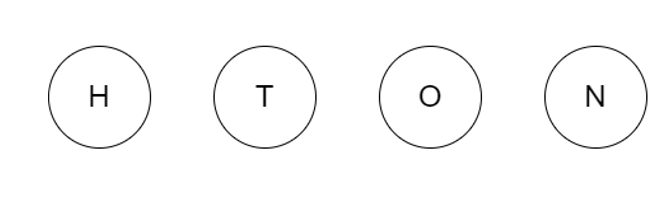


Рисунок 1 – Узлы для ключа «HTON»

Число 9 в качестве ключа будет предварительно переведено в двоичную систему и разбито на биты 1001 как показано на рисунке X.

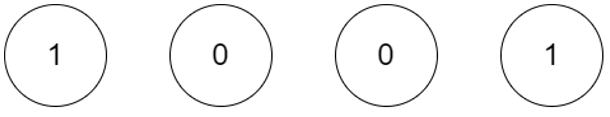


Рисунок 2 – Узлы для ключа 9

После разделения ключа, необходимо поочерёдно вставить узлы в дерево. Вставка всегда происходит из корня. Алгоритм вставки может быть описан так: пусть нам необходимо вставить ключ k, который разбивается на , … ... узлы, мы находимся в узле дерева и уже вставили все узлы до .

Тогда,

* если уже имеет ребёнка (назовём его ), то мы пропускаем вставку (так, как он уже есть = ) и вставляем начиная c .
* если не имеет ребёнка , то мы добавляем в ребёнка (назовём его ) и вставляет начиная с .

В обоих случаях, если был обработан последний узел (), то помечается как терминальный. Это указывает, что данный узел является в дереве концом какого-то ключа.

Значения, ассоциированные с ключом, привязываются к терминальным узлам. Таким образом, не каждый узел в префиксном дереве связан с определённым значением.

Пример такого дерева продемонстрирован на рисунке X. В нём префиксное дерево, включает в себя строки «KAFKA», «KARMA», «KAFKAEAESQUE», «JAIL». Узлы красного цвета являются терминальными, белого – промежуточными, зелёный узел соответствует корню.

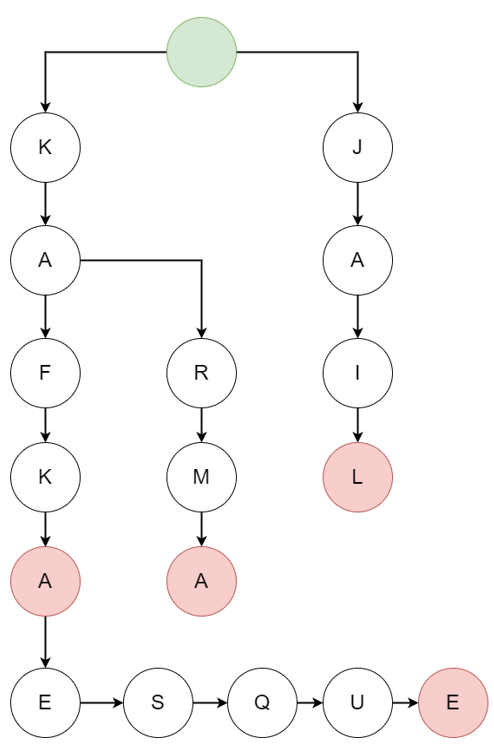


Рисунок 3 – Префиксное дерево, хранящее строки

Стоит заметить, префиксное дерево компактно хранит ключи: узлы строк с одинаковыми префиксами не дублируются, а хранятся в единственном экземпляре («KAFKA» и «KARMA» имеют общий префикс «KA»). Это и делает эту структуру данных эффективной по памяти. То же можно сказать и про Hash Array Mapped Trie, которая использует свойства префиксных деревьев.

Префиксное дерево может быть реализовано несколькими способами. Первый способ – «наивный». Пусть у нас есть определённый алфавит, состоящий из символов (обычно = 256, что соответствует размерности таблицы ASCII). То есть, всего может быть различных видов узлов в дереве, и у каждого узла может быть не больше, чем различных детей. Тогда каждый узел хранит ссылки на своих детей в собственном фиксированном массиве на элементов (). Если , значит данный узел не имеет ребёнка, узла с типом .

На рисунке X изображён узел A, который имеет только трёх детей B, D, E но хранит информацию об отсутствии 253 других возможных детей (например, , то есть отсутствует узел C).

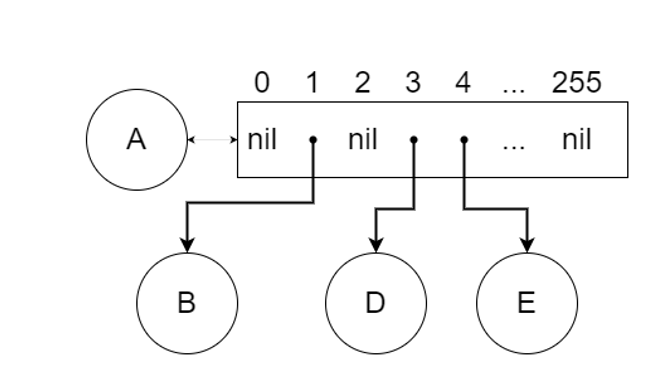


Рисунок 4 – Узел в наивной реализации префиксного дерева

Данный подход имеет плюс в скорости поиска следующего узла (асимптотика – Вместе с тем, огромное количество памяти расходуется впустую на хранение .

Закономерным решением проблемы с памятью было бы использование списков вместо массивов. Это увеличит время операции до , где m – количество детей у узла, но позволит избавиться от лишнего расхода памяти. На рисунке X показано префиксное дерево, реализующее   
данный подход: узел A хранит связный список из трёх элементов B, D, E, что разы меньше, чем в наивной реализации.

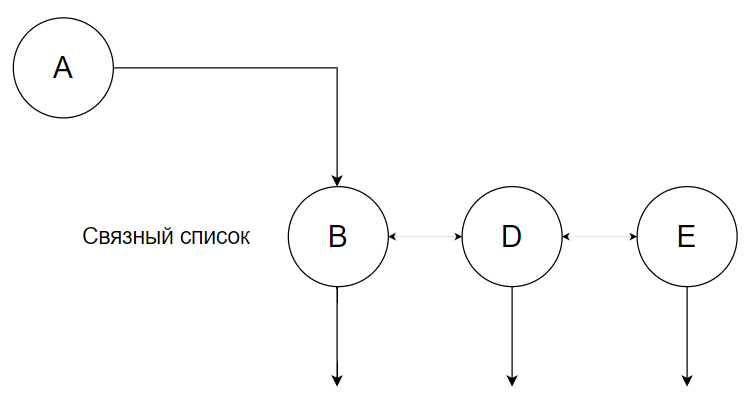


Рисунок 5 – Узел в реализации префиксного дерева со связным списком

Hash Array Mapped Trie снова находится на середине между двумя подходами. Для хранения детей использует массив как в наивном методе (но не статический, а динамический). И так же, как и реализации со связным списком, не хранит лишние элементы.

Основные положения Trie:

* Позволяет поддерживать словарь (операции: вставка, удаление, поиск)
* Ключ делится на части, которые репрезентируются в узлы.
* Узлы бывают трёх видов: терминальные (хранят в себе значение), корень (представляет особый узел, у которого нет родителей), промежуточные
* При вставке узлы ключа последовательно добавляются в дерево. Если два ключа имеют общий, префикс, их узлы, соответствующие этому префиксу, не дублируются.
* Каждый узел хранит детей в связном списке, что позволяет экономить память (в отличии от статического массива)

1.2 **Описание однопоточной Hash Array Mapped Trie**

После детального рассмотрения префиксного дерева, можно приступить к описанию устройства Hash Array Mapped Trie [9]. Hash Array Mapped Trie (HAMT) представляет собой такое же дерево поиска для хранения, удаления, поиска ключей/значений. Как было сказано ранее, она сочетает в себе свойства хеш-таблиц и префиксных деревьев. Но также включает другие идеи для повышения быстродействия.

Как и для префиксных деревьев, HAMT разделяет свой ключ на части. Эти части определяют путь до ключа/значения. Ключи могу иметь общие префиксы, что позволяет экономить память. Но, если в префиксных деревьях разбиение определялось для каждого типа ключа по-разному (для строки – на символы, для числа – на биты), то Hash Array Mapped Trie разбивает все типы ключей единообразно: генерирует 64-битный хеш ключа, а затем делит его части по пять битов. Итого, частей . Каждые пять бит – это отдельное число. Например, если в дерево добавляется ключ, который хешируется в число 2172 (двоичную форму будем обозначать так ), то он будет разбит на

Части репрезентируются в узлах 31, 16, 0, 0 и так далее (рисунок X).

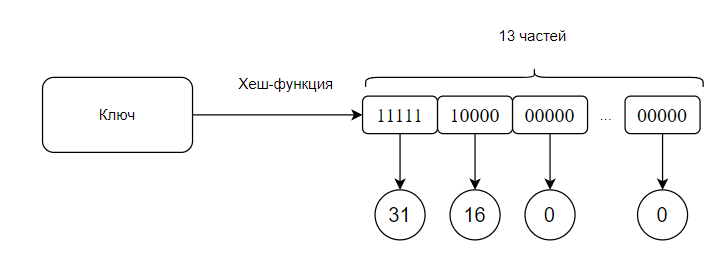


Рисунок 6 – Разбиение хеша ключа на части

Таким образом, данный ключ в дереве будет располагаться по пути, который показан на рисунке X. Зелёным цветом обозначен корень дерева, белым – промежуточные вершины, красным – терминальная вершина, где хранится значение.

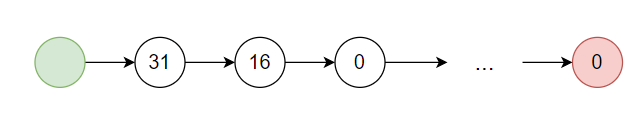


Рисунок 7 – Путь в дерева для ключа с хешом 2172

На самом деле, для хеша 2172 не обязательно хранить все узлы в контексте дерева. Нужно хранить ровно такое количество, чтобы можно было уникальное идентифицировать путь до ключа в дереве. Приведём несколько примеров.

Пусть у нас есть пустое дерево (рисунок X)



Рисунок 8 – Hash Array Mapped Trie, не содержащий ни одного ключа

Добавим в дерево ключ с хешом 2172 (рисунок X). Как говорилось ранее, нужно добавлять минимальное количество узлов, чтобы можно было уникально идентифицировать путь до ключа. В данном случае в пустом дереве хватит лишь одного узла. Этот узел нужно сделать терминальным и положить в него значение и полный ключ (это понадобиться для проверки при поиске).

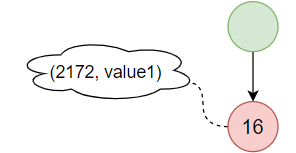


Рисунок 9 – Hash Array Mapped Trie c единственным ключом

Теперь, если нам необходимо будет найти ключ 2172, мы будем проходить не по всем узлам (ведь дерево их не содержит), а лишь по части из них. И когда поиск приведёт к терминальной вершине, необходимо провести проверку на полное совпадение ключей. Ниже будет описан более строгий алгоритм поиска.

При добавлении ключа с хешом 63 появляется неоднозначность (рисунок X). Действительно, у обоих хешей первая часть совпадает. Значит, узел 31 () уже нельзя сделать терминальным (неоднозначно какому ключу он будет принадлежать: 63 или 2172). Поэтому необходимо сделать узел 31 промежуточным, и отвести от него узлы 16, 1 и сделать их терминальными. Эти узлы вычисляются уже по второй части хеша.

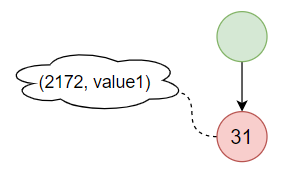


Рисунок 10 – HAMTs c двум ключами с общим префиксом у хешей

Может случиться так, что и вторые части хеша будут совпадать. Тогда производим те же самые операции (создаём общий узел и отводим узлы для детей). В случае, если все части хешей совпадают, то на последнем уровне вместо терминальной вершины создаём *терминальный узел-список* из двух ключей. Терминальный узел-список – это список, который хранит в себе множество пар ключ/значение.

На рисунке X показан процесс разрешения коллизий при добавлении двух ключей с идентичными первыми 12 частями хеша.

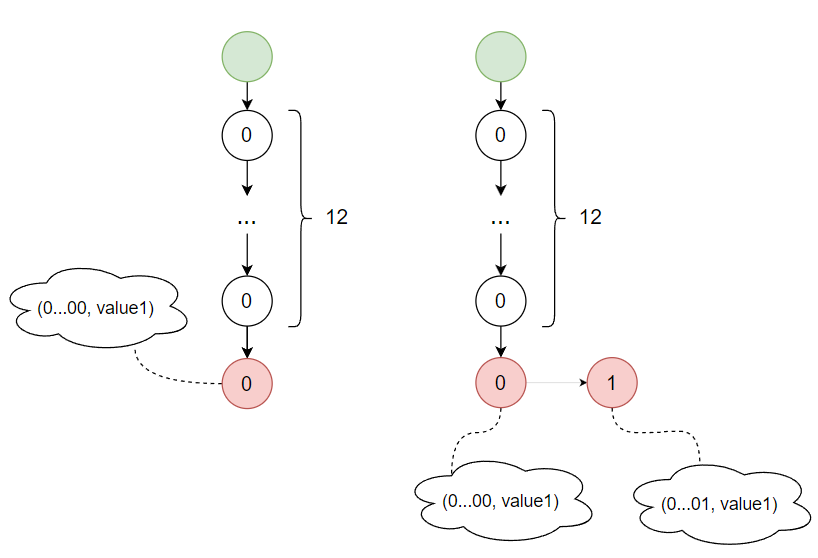


Рисунок 11 – Разрешение коллизии

Теперь осуществим удаление. Оно будет производиться рекурсивно, начиная с терминальной вершины, соответствующей ключу: если текущая вершина не имеет детей, то она подлежит удалению. То есть, HAMT сжимает дерево сразу же после удаления, поэтому не хранит лишних узлов.

Более строго операции вставки, поиска, удаления можно описать следующим образом.

*Вставка*: пусть нам необходимо вставить пару , где , … ... , мы находимся в узле дерева и уже вставили все узлы до . Тогда:

* если у текущего узла дерева нет ребёнка , мы добавляем его, помечаем терминальным и вкладываем
* если у текущего узла есть нетерминальный ребёнок , то пропускаем вставку и производим добавление следующей части хеша , но уже c узла
* если (то есть, мы находимся ниже предпоследнего уровня глубины дерева) и у текущего узла есть терминальный ребёнок с ключом и значением , то сначала проверяем не совпадает ли он с . Если да, то заменяем внутри этого ребёнка на. Иначе, делаем из него промежуточный узел и добавляем , и
* если (то есть, мы находимся на предпоследнем уровне глубины дерева) и у текущего узла есть терминальный ребёнок с ключом и значением , то мы заменяем его на терминальный узел-список, состоящий из
* если (то есть, мы находимся на предпоследнем уровне глубины дерева) и у текущего узла есть терминальный узел-список ребёнок, то проверяем есть ли в этом списке пара , где . Если имеется, то заменяем на . Иначе, добавляем в список новую

*Поиск:* пусть нам необходимо найти ключ , где , … ... , мы находимся в узле дерева и уже спустились на уровень . Тогда:

* если у текущего узла дерева нет ребёнка , то ключ/значение не найдены.
* если у текущего узла есть нетерминальный ребёнок , то производим поиск следующей части хеша , но уже с узла .
* если у текущего узла есть терминальный ребёнок с ключом и значением, то проверяем есть ли в этом списке пара , где . Если да, то и есть искомая пара. Иначе, ключ/значения не найдены.

*Удаление*: пусть нам необходимо удалить ключ . Произведём поиск этого ключа. Если ключ отсутствует, то завершаем удаление. Иначе, существует терминальная вершина соответствующая ему. Начнём удаление с вершины . Тогда:

* если текущая промежуточная вершина не имеет детей, то удаляем её (удаляем ссылку на неё у родителя)
* если текущая терминальная вершина не имеет детей, то удаляем её, если она хранит единственный ключ. Если это узел-список, то удаляем из неё такую пару , что .
* иначе, прекращаем процесс.

Следующий важный момент в Hash Array Mapped Trie – это использование связки битмапа и динамический массив. Идея заключается в том, что узел хранит ссылки только на существующих детей в динамическом массиве. Ранее поиск по такому массиву занимал *O(n) (*как это было в одной из реализаций trie ранее). Но HAMT также хранит битмапу, по которой можно определить индекс необходимого элемента в массиве за время . Индекс вычисляется с использованием побитовых операций, что значительно быстрее предыдущей версии.

Структура данных битмапа (bitmap) – это число с определённой разрядностью , где каждый бит говорит об присутствии (если равен единице) или отсутствии (если равен нулю) элемента в связанном с ней объектом [10]. В случае Hash Array Mapped Trie битмапа есть у каждого узла, её разрядность 32, а характеризует она наличие или отсутствие детей в динамическом массиве. Например, битмапа 104 (или в двоичном виде ) у узла P (рисунок X) говорит о наличии трёх детей: 0, 1, 3.

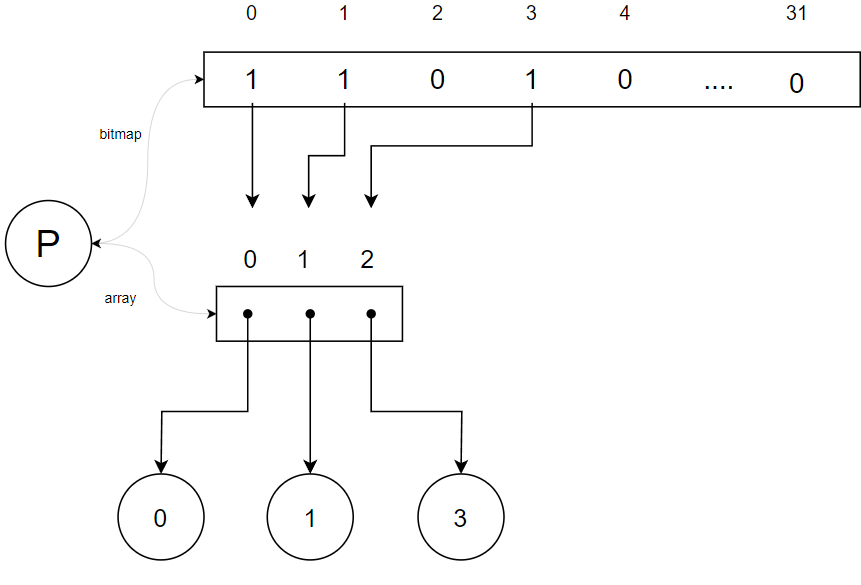


Рисунок 12 – Узел с битмапой и массивом (array), содержащий трёх детей

Зная индекс битмапы, где поднят флаг единицы, можно вычислить индекс в массиве.

Формула искомого индекса вычисляется по формуле (1):

, (1)

где – функция подсчёта поднятых битов в числе [11], – позиция в битмапе с установленным битом, & - операция побитового «И», – побитовый сдвиг числа 1 раз влево. Получения бита в битмапе производится тоже через побитовые операции: . Итого, чтобы получить элемент массива по битмапе, необходимо сделать два шага:

1. Проверить, установлен ли бит на определённой позиции (используются только побитовые операции)
2. Если в предыдущем шаге бит не установлен, возвратить ничего. Иначе, вычисляем индекс массива (используются только побитовые операции)

Благодаря битмапе поиск ребёнка в узле происходит не за время , а за .

Основные положения о Hash Array Mapped Trie, описанные в этой главе:

* Ключ хешируется и разбивается на части.
* У каждого узла может быть не больше, чем 32 ребёнка
* Максимальная высота дерева равна 13
* При добавлении ключей с одинаковыми префиксами хешей, общая часть не дублируется.
* Добавляется минимальное количество узлов, при котором можно уникально идентифицировать ключ
* При удалении, дерево сжимается и не хранит лишних узлов.
* Каждый узел хранит детей в динамическом массиве и битмапе. Через побитовые операции можно за время вычисляет ребёнок узла или его отсутствие.

1.3 **Описание неблокирующей Hash Array Mapped Trie**

Идея алгоритма [8] построена на введении трёх типов узлов (рисунок X):

* SNode – узел, который содержит список пар ключ/значение
* CNode – узел с битмапой и массивом детей. Дети могу двух типов: SNode и INode
* INode – узел, который хранит в себе ссылку на CNode. Это искусственная обёртка над CNode для поддержания неблокирующей реализации.

Все три типа унаследованы от абстрактного узла Node.

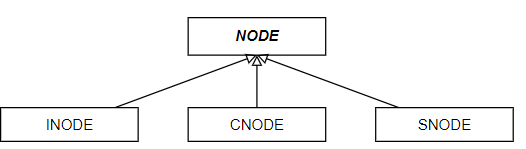


Рисунок 13 – UML–диаграмма для узлов

Начнём описание с SNode и CNode. В дальнейших рассуждениях будем усложнять структуру данных.

Пусть у нас есть дерево с узлом SNode (рисунок X).

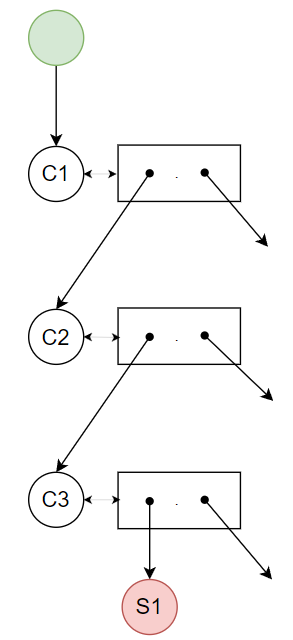


Рисунок 14 – Hash Array Mapped Trie с узлом S1

Тогда, если мы хотим вставить в С3 ещё одного ребёнка, нам необходимо будет совершить два действия:

1. Обновить битмапу, чтобы она указывала и на новый элемент
2. Вставить новый узел в массив C3

Операции 1 и 2 происходят не атомарно и это приводит к логической гонке данных. Например, если два потока T1 и T2 одновременно пытаются добавить соответственно новые узлы S2 и S3, то может произойти следующее:

1. T1 копирует битмапу и меняет её локально.
2. T2 копирует битмапу (которая указывает только на S1) меняет её локально.
3. T1 меняет битмапу в C3 на изменённую локальную версию
4. T2 перезаписывает битмапу C3 на свою изменённую локальную версию.

В результате, битмапа будет содержать версию потока T2, который не учитывает изменения T1. Решение этой проблемы заключается в том, что обновляют не битмапу и массив, а целые узлы. Так, в данном случае T1 не меняет содержимое C3, а копирует этот узел , меняет его локально, и уже потом обновляет в С2 указатель на С3 на

Для того, чтобы избежать техническую гонку данных при изменении указателя, необходимо воспользоваться операцией *compare-and-swap*   
(cas) [12].

Это функция принимает три аргумента:

1. указатель на старый узел (назовём это )
2. значение, которое лежало по этому указателю (назовём это )
3. обновлённый узел (назовём это )

Если , то мы меняем содержимое указателя на и возвращаем , как показатель того, что значение было обновлено. Иначе возвращаем .

В данном случае, поток T1 сначала зафиксирует локально чему равно С3, сделает его копию и начнёт менять (добавит новый узел). Далее воспользуется сas, куда передаст указатель на C3 в качестве p, зафиксированный ранее в качестве и в качестве . Если cas прошёл успешно – это означает, что никакой другой поток не поменял C3 и наш поток обновил его. Иначе, производим вставку заново с самого начала. Операция cas происходит за одну ассемблерную инструкцию, что позволяет считать её неблокирующей.

Даже при таком решении с атомарным compare-and-swap над узлом, остаются проблемы с гонками данных. Покажем это на примере. Пусть есть некоторое дерево в следующей конфигурации (рисунок X) и пусть есть поток T1, который хочет вставить в C3 новый узел S2. Для этого он копирует узел С3 (жёлтым цветом обозначена копия).

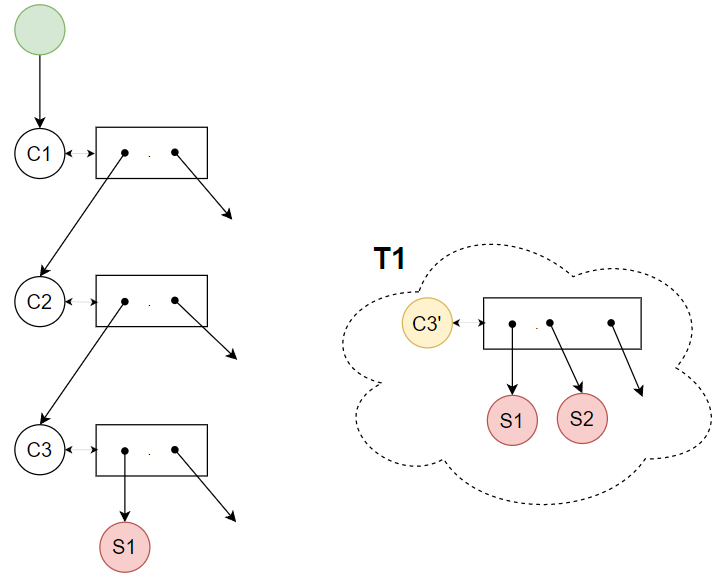


Рисунок 15 – поток T1 копирует C3 и добавляет в копию новый узел S2

Пока T1 занимается копирование и локальным добавлением нового узла, другой поток T2 решает добавить новый узел S3 в C2.

Для этого он также создаёт копию , добавляет узел S2 (рисунок X).

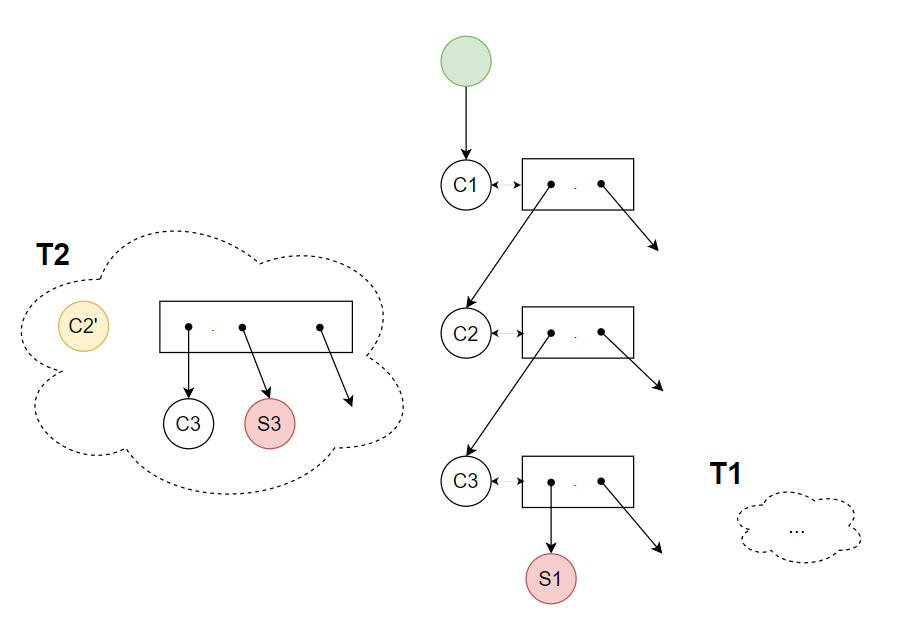


Рисунок 16 – поток T1 копирует C3 и добавляет в копию новый узел S3

И если сначала T1 сделает cas-операцию, а потом T2, то узел S2 потеряется (рисунок X).

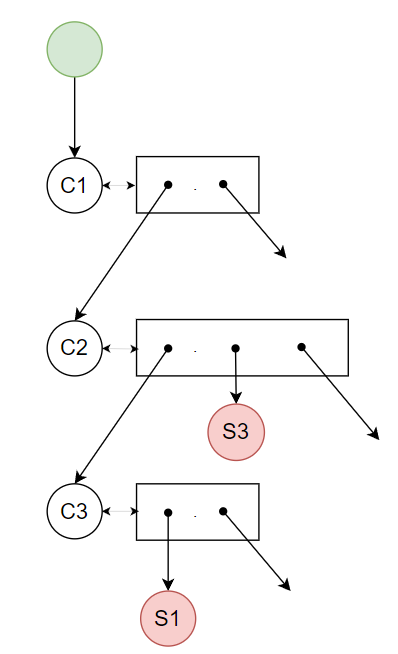


Рисунок 17 – дерево после обновление C3 потоком T1 и С2 потоком T1

Для решения данной проблемы, вводится новый тип узла INode. Он является обёрткой над каждым CNode. То есть, для каждого CNode узла существует родитель INode, который ссылается только на него. Тогда CNode в качестве детей может иметь INode или SNode узлы. Теперь любая cas-операция производится над содержимым промежуточного узла INode, что позволяет избавить от проблемы, описанной ранее.

На данном этапе, неблокирующая Hash Array Mapped Trie работает корректно для операций вставка и поиск. Но есть ещё операция удаления. Она происходит следующим образом: удаляется пара с искомым ключом из соответствующего узла. Если в этом узле SNode не осталось ключей, то необходимо удалить его и произвести сокращение дерева. Сокращение дерева случается, когда существует INode, ребёнок которого имеет одного ребёнка. Тогда этот ребёнок должен заменить INode. Во время этого процесса может возникнуть проблема. Представим дерево, которое, должно сократиться (рисунок X) потоком T1. Пока T1 копирует и изменяет C1(меняет дочерний элемент I2 на S2), другой поток T2 добавляет в C2 новое значение S3. Для этого также начнёт копировать C2 и обновлять копию. В этот момент, T1 совершает сокращение (I2 заменяется на S2). Затем поток T1 совершит cas и изменит дочерний элемент I2. Но I2 уже удалён. Поэтому, в дереве не будет вершины S3.

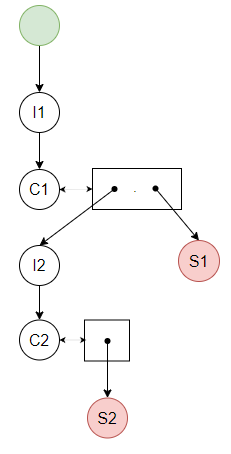


Рисунок 18 – дерево, перед тем, как сократиться

Решение этой проблемы – это помечать сокращаемые узлы CNode как гробовые (tomb-node). Их все потоки могут только сокращать. Если CNode уже нельзя сократить, то необходимо его воскресить (из гробового узла сделать обычный).

Основные положения неблокирующей Hash Array Mapped Trie, описанные в этой главе:

* вводятся три типа узлов INode, CNode, SNode
* при помощи compare-and-swap атомарно происходит изменение целого узла, а не битмапы и массива отдельно
* после удаления узла, дерево начинает сжиматься
* CNode, который может быть сокращён помечается как гробовой. В такой узел нельзя добавлять или удалять дочерние элементы.

# 

# 2 РАЗРАБОТКА HASH ARRAY MAPPED TRIE

В предыдущей главе описаны теоретические аспекты алгоритма неблокирующего алгоритма Hash Array Mapped Trie и также заложены предпосылки для внедрения. Внедрение такого алгоритма — процесс трудозатратный. Он включает в себя множество пунктов: постоянная отладка программы, ручное и автоматическое тестирование, анализ кода, проверка гипотез, исправление ошибок и так далее. Чтобы облегчить эти процессы, необходимо иметь чёткий план действий: что реализовывать в первую очередь, а что можно отложить на потом.

План:

* Настройка рабочего окружения
* Реализация однопоточной версии HAMT
* Реализация блокирующей многопоточной версии (с грубой и тонкой синхронизацией)
* Переход к неблокирующей многопоточной версии HAMT
* Решение проблемы ABA
* Внедрение в библиотеку libcds

2.1 Настройка рабочего окружения  
 В качестве контроля версий был выбран Git [13]. Он является свободно распространяемым, наиболее популярным и имеет необходимый функционал:

* создание веток разработки
* слияние одной ветки в другую
* просмотр истории файла
* создание именованных коммитов
* удобное взаимодействие через командную строку
* множество интеграций с другими сервисами

Для того, чтобы код не хранился в одном месте, а также была возможность поделиться с ним, использовался GitHub [14]. Это бесплатный сервис, который работает с системой контроля версий Git и позволяет хранить, анализировать код, выполнять Git команды, совместно редактировать код и тестировать его благодаря GitHub Actions [15]. С использование GitHub Actions можно настроить запуск тестирования кода при каждом коммите. В основном репозитории расположен файл Dockerfile (рисунок X), который запускает докер контейнер при коммите [16]. В этом файле описывается установка операционной системы, копирование кода из текущего репозитория GitHub и запуск скрипта entrypoint.sh.



Рисунок 19 – Код файла Docker

В скрипте entrypoints.sh (рисунок X) производится подготовка для сборки и запуска исполняемого файла для тестирования.

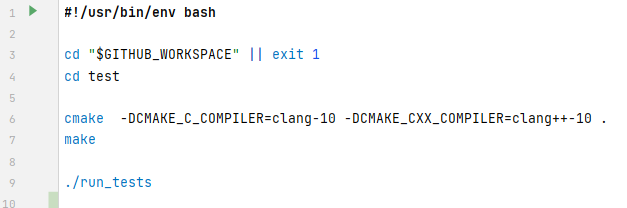


Рисунок 20 – Код файла entrypoints.sh

Для сборки был написан CmakeFileLists.txt. Этот файл используется утилитой CMake [17], которая генерирует Makefile. А Makefile, в свою очередь, используется утилитой Make [18], которая создаёт на его основе исполняемый файл для тестирования.

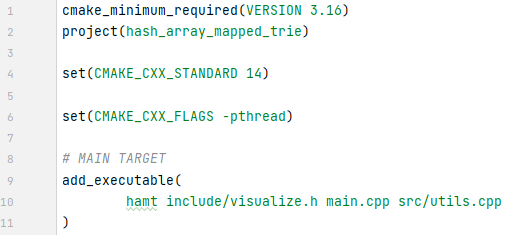


Рисунок 21 – Код файла CmakeFileLists.txt

Для сборки был выбран СMake. В качестве интегрированной среды разработки использовался Clion от компании JetBrains [19]. Помимо встроенного отладчика, статических анализаторов кода, Clion предоставляет удобный графический интерфейс к Git, а также интегрируется с GitHub. Одна из особенностей неблокирующих алгоритмов – это то, воспроизведение ошибок в код не постоянно. Для увеличения вероятности воспроизведения потенциальных ошибок в код тестирование запускалось по расписанию при помощи GitHub Actions. Так, в специальный файл, отвечающий за запуск докер контейнера, были добавлена следующие строки (рисунок X):



Рисунок 22 - Файл с конфигурацией запуска тестов

В поле cron с использование синтаксиса подобного утилиты crontab находится указание на запуск контейнера с тестами по расписанию «каждые пять минут». В случае, если тесты проходят неуспешно, на почту отправляется соответствующее уведомление. Автоматическое тестирование при коммитах и по расписанию облегчает внесения изменений в код. Также это ускоряет разработку, так как избавляет от необходимости ручного тестирования и ожидания результатов.

Ещё одна проблема при разработке неблокирующих структур данных – сложность отладки. Интегрированная среда разработки Clion предоставляет возможности для этого, но таких функций недостаточно. Зачастую необходима визуализация самой структуры данных Hash Array Mapped Trie. Для этого был добавлен код для визуализации (visualize.h), в котором была реализована логика сбора информации о переданном HAMT, её обработка, а затем передача библиотеке boost graphviz [20]. Данная библиотека умеет визуализировать графы в формате в png. На рисунке X показан пример визуализации произвольного Hash Array Mapped Trie. Узел I0 является INode и имеет CNode ребёнка C1. В листьях дерева содержаться SNode. Так узел C6 хранит ссылку на дочерний элемент SNode с ключом 8 и значением 32.

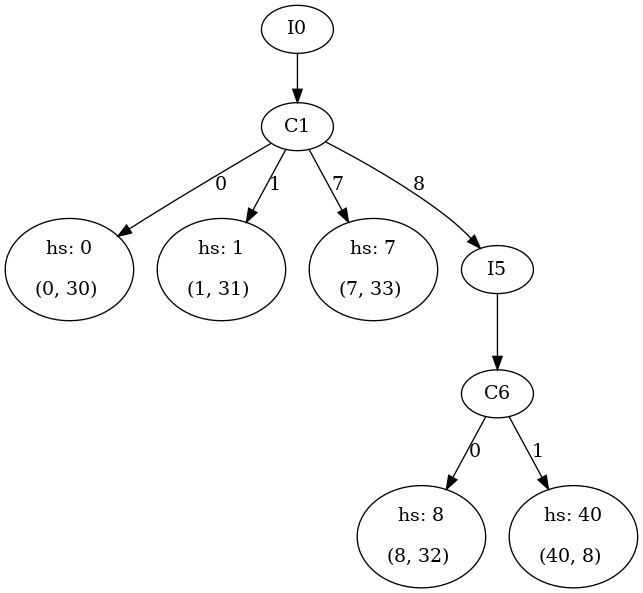


Рисунок 23 – Дерево, построенное при помощи visualize.h

2.2 Реализация однопоточной версии

Приступая к описанию однопоточной версии HAMT на языке C++, необходимо сказать, что с изначально код разрабатывался с учётом будущей трансформации под lock-free версию HAMT. Это было сделано для ускорения разработки и логического структурирования всей работы.

Начнём с перечисления сущностей (классов, структур), опишем их взаимодействия друг с другом. А затем перейдём к трём основным операциям insert, remove, lookup.

Есть класс Hamt, представляющий контейнер c Hash Array Mapped Trie структурой (рисунок X). В нём есть поле root, в котором хранится корневой узел и три публичных метода: вставка, удаление и поиск. Тип K обозначает один из произвольных базовых типов в С++ (int, string, float …).

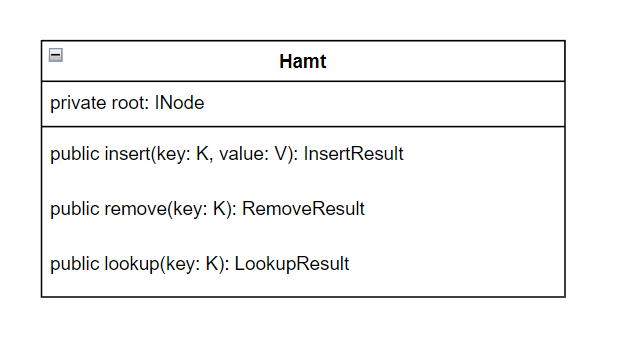


Рисунок 24 – UML-диаграмма для Hamt класса

При создании контейнера, необходимо указать какого типа буду ключи и значения (рисунок X)



Рисунок 25 – Создание контейнера с ключами и значениями типа int

Все три метода insert, remove, lookup возвращают соответственно структуры InsertResult, RemoveResult, LookupResult для хранения дополнительной информации вставке, поиске и удалении.

Например, на рисунке X показана UML-диаграмма для структуры RemoveResult. Тип Status – тип перечисления. Он может хранить в себе три значения:

* Removed (если удаление прошло успешно, тогда в поле value будет храниться значение у удаляемой пары)
* Failed (заготовка для неблокирующей версии, где удаление может не получиться)
* NotFound (в дереве нет такого ключа, которого хотят удалить)

В value хранится значение, которое было удалено.

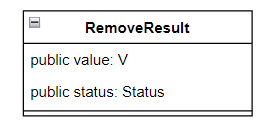


Рисунок 26 – UML-диаграмма для узлов

Внутри дерева есть три вида узлов, наследующиеся от базового абстрактного типа Node (рисунок X).

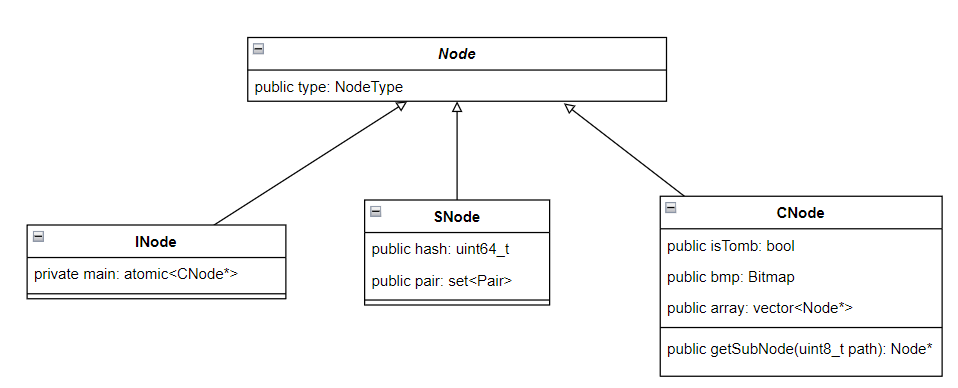


Рисунок 27 – UML-диаграмма для узлов

INode, как говорилось в прошлой главе, является обёрткой на CNode. Это позволяет обеспечить потокобезопасное обновление узлов. Именно поэтому у него есть одно поле main типа [21]. С таким типом можно совершать атомарные операции над указателем на CNode (в том числе и cas-операцию). Сейчас это лишь заготовка для будущей lock-free версии, и она не противоречит однопоточной.

SNode – узел состоящий из двух полей: хеш и множество пар ключ/значение с таким хешом. Поле hash имеет тип uint64\_t, что занимает 64 бита. Ровно столько, чтобы вместить в себя максимальное значение хеша. Это позволяет не хранить лишние биты и не тратить память. Pair представляет собой класс с двумя публичными полями (один для ключа другой для значения) и перегруженным оператором меньше (это необходимо для создания множества уникальных элементов на его основе).

CNode – хранит в себе поле isTomb (заготовка для lock-free версии, указывает «гробовой» это узел или обычный), поле типа Bitmap (битовая карта у данного узла), и динамический массив указателей на узлы типа Node. Один из важнейших методов – getSubNode, который принимает тип ребёнка (число от 0 до 31) и возвращает указатель на ребёнка или nil, если такого ребёнка нет.

Подробнее стоит остановиться на вспомогательной структуре данных Bitmap и методах, работающих с ней (рисунок X). Bitmap характеризует состояние массива array и указывает на отсутствие или наличие определённых детей. При удалении или добавлении новых узлов в массив, поле типа Bitmap у CNode должно изменяться. Для у это у Bitmap есть публичные методы set, unset и isSet.

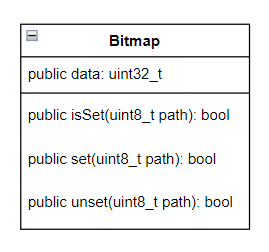


Рисунок 28 – UML-диаграмма для класса Bitmap

Все три метода (рисунок X) используют побитовые операции «И» (&), сдвига вправо (<<), инвертирование (~) [22].

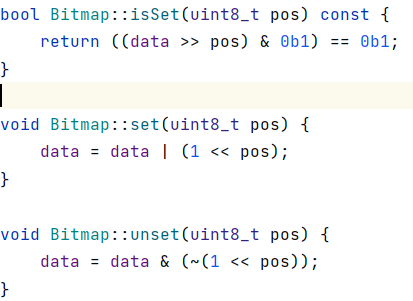


Рисунок 29 – Реализация методов Bitmap

На основе данных битмапы также вычисляет индекс массива (рисунок X). Эта логика реализована в отдельном методе класса CNode.

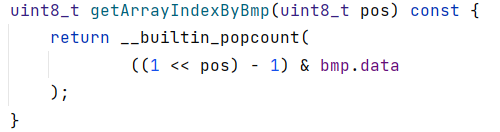


Рисунок 30 – Вычисление индекса массива для установленного бита

Команда вычисляет количество установленных битов в числе [23], – значение битмапы в соответствующем узле CNode.

Теперь, когда были описаны основные классы и структуры, можно перейти к реализации insert, remove, lookup операций. Они определены в публичных методах класса Hamt и каждый ссылается на соответствующий непубличный метод. Так было сделано по двум причинам:

* Декомпозиция метода (в публичном создаются вспомогательные структуры, которые идут в приватные)
* Необходимость повторного вызова вставки, удаления и поиск (это заготовка для lock-free версии)
* Такая структура легко трансформируется в блокирующую многопоточную версию.

Для примера разберём операцию insert (рисунок X). На вход в публичную функцию поступают ключ и значение. В строчках 282-284 идёт проверка особого случая (проверка на пустоту дерева). В строчке 285 создаётся новый узел SNode, который необходимо вставить. Далее, вызывается приватный метод insert (рисунок X).

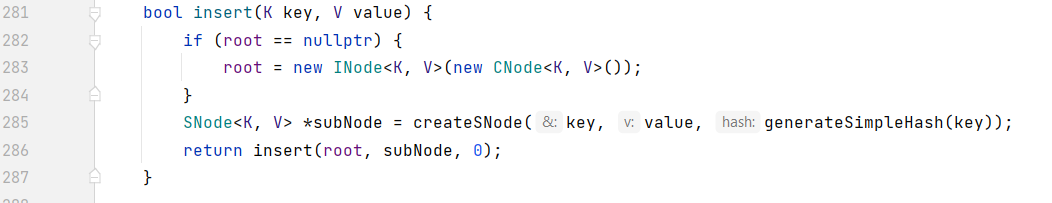


Рисунок 31 – Вычисление индекса массива для установленного

В приватном методе, чтобы определить по какому пути идти, вычисляется часть хеша (строчка 347). Далее вычисляется дочерний элемент, на которого мы, возможно перейдём на следующем шаге. И затем обрабатывается одна из следующих ситуаций, которая была описана в предыдущей главе.

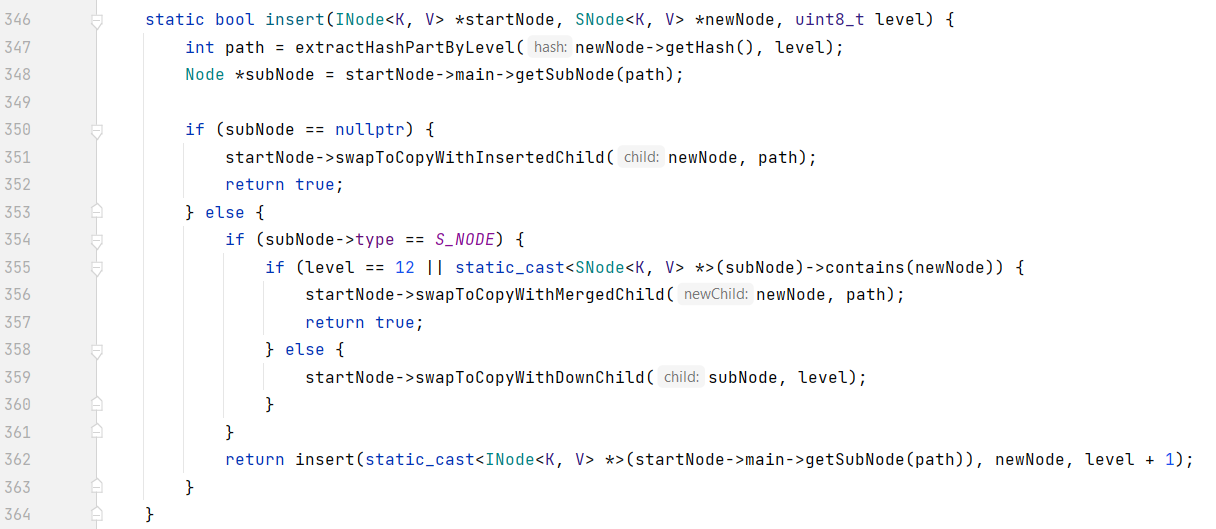


Рисунок 32 – Вычисление индекса массива для установленного бита

Если по данному пути нет ребёнка, то есть (строка 350), то для текущего узла вызывается специальный метод swapToCopyWithInsertedChild, который вставляет ребёнка по пути path в текущий узел (строка 351).

Методы , , занимаются обработкой одной из возможных ситуаций. Все из перечисленных изменяют состояние текущего узла.

2.3 Реализация блокирующей многопоточной версии

Код, описанный в предыдущей главе, легко трансформируется в многопоточную блокирующую версию. Всё что, необходимо поменять – это добавить примитив синхронизации в класс Hamt (рисунок X) и определить, в каких местах он будет использоваться (то есть определить где будут критические секции).

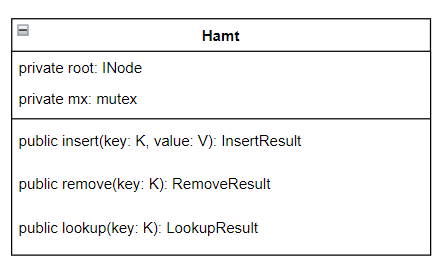


Рисунок 33 – UML-диаграмма класса Hamt для блокирующе многопоточной версии

В качестве примитива синхронизации был выбран мьютекс. Чтобы попытаться захватить мьютекс (начало критической секции), потоку необходимо воспользоваться методом . Если это удаётся сделать, то он индивидуально входит в критическую секцию. Иначе, он ожидает. После завершения работы в критической секции потоку необходимо освободить примитив при помощи метода .

Как говорилось ранее, методы вставки, удаления, поиска логически разделены на публичные и приватные. Такая структура позволяет легко добавить использование мьютекса. Идея в том, что потоконебезопасный код внутри приватных методов необходимо окружить взятием и отпускание мьютекса (рисунок 2.3.2). Такая реализация будет считаться грубой, так как критические секции охватывают весь метод.

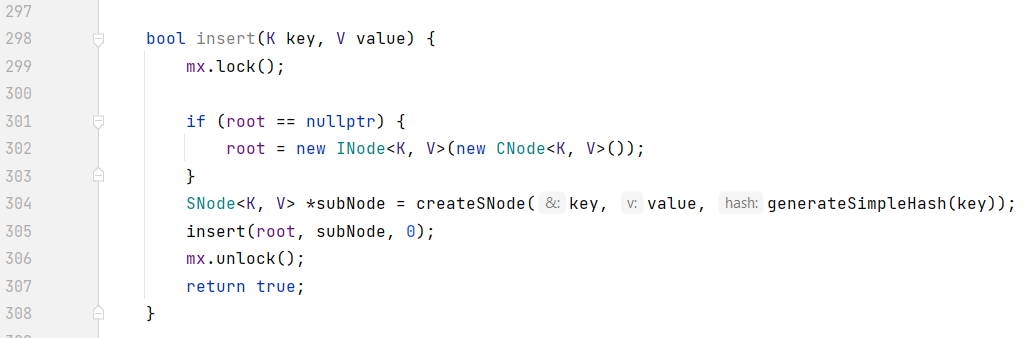


Рисунок 34 – Реализация публичного метода вставки в блокирующей версии

Для тонкой синхронизации блокирующей версии, необходимо добавить мьютекс на каждый узел INode. Когда поток доходит до соответствующего узла, он должен попытаться захватить примитив синхронизации. После завершения операции или перехода на следующий узел, необходимо отпустить мьютекс. Таким образом, при данном подходе потоки могут работать одновременно в разных частях дерева

2.4 Реализация неблокирующей версии

Неблокирующая версия HAMT является закономерным продолжением блокирующей многопоточной версии, которая, в свою очередь, была трансформирована из однопоточной. Поэтому при реализации lock-free версии необходимо было постепенно сужать критические секции в блокирующей версии и вводить cas-операции.

Главным отличием неблокирующей реалиации является то, что узлы дерева изменяются только при cas операции. Поток всегда сначала копирует узел, который он хочет изменить, меняет эту локальную копию и уже потом пытается вставить свою версию. Если вставка своей версии (операция cas) проходит неуспешно, значит другой поток уже заменил соответствующий узел и наши изменения неактуальны. Поэтому необходимо повторить операцию.

Код для вставки проходит через следующие этапы (рисунки X и X):

1. Запускается публичный метод, в бесконечном цикле происходит попытка вставить новый узел. Либо вставка происходит в корне, либо вызывается приватный метод insert для вставки ниже корня.
2. В приватном методе фиксируется текущая CNode и родитель, если он есть. (строки 398-399)
3. Попытка сократить дерево (строки 441-443). Метод производит сокращение текущего узла, если в этом есть необходимость. Возвращаемое значение сигнализирует о том, необходимо ли заново провести операцию вставки или нет.
4. Получение копии текущего CNode (строка 445)
5. В зависимости от дочернего элемента, произвести изменения локальной копии (строки 450-475)
6. Попытаться вставить новую версию CNode при помощи cas-операции (строки 453, 463, 467, 453)
7. Если не получилось, то повторяем вставку через вызов публичного метода (строка 474)

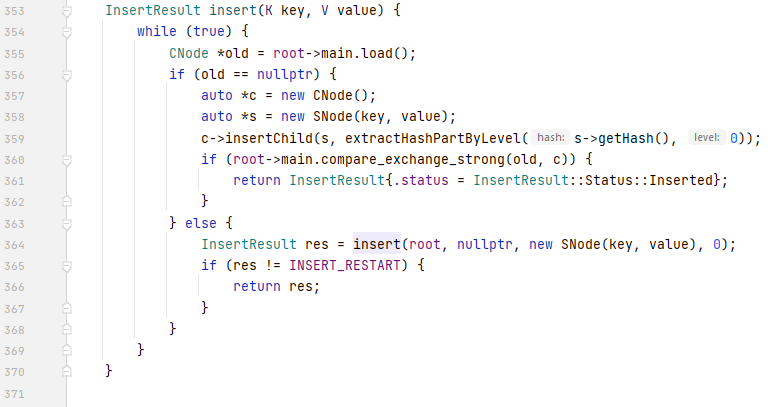


Рисунок 35 – Реализация приватного метода вставки в lock-free версии

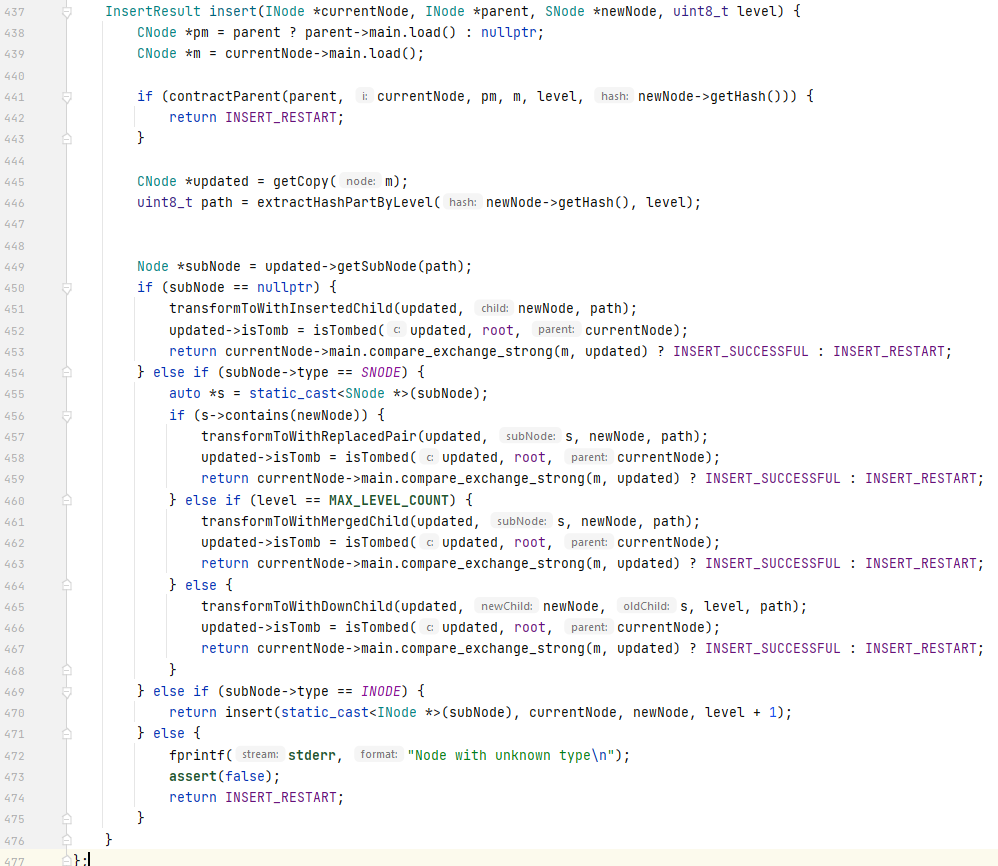


Рисунок 36 – Реализация приватного метода вставки в lock-free версии

В данной реализации было отдано предпочтение нерекурсивному алгоритму. Это помогает избежать проблемы переполнения стека при большом количестве неуспешных попыток произвести вставку, а значит большом количестве рекурсивных вызовов функций.

2.5 Решение проблемы ABA в неблокирующей версии

Проблема ABA возникает [24], когда поток считывает из ячейки данных дважды значение, это значение оказывается одинаковым и это воспринимается как «ячейка данных не менялась». Однако, это может быть неверным, ведь между двумя считываниями другой поток мог поменять указатель, изменить значение по нему, а потом вернуть прежний указатель.

Существует несколько подходов для решения проблемы ABA. Один из них hazard pointers («опасные указатели»). Данная схема позволяет освобождать ячейку памяти только когда на неё никто не ссылается.

«Опасные указатели» реализованы в библиотеке libcds [25]. Это позволяет внедрить их в код и избавиться от ABA. Для этого любые места памяти, которые необходимо удалить, помечаются как «опасные указатели». Такие указатели удаляются отложено, когда это можно будет сделать корректно.

Чтобы подключить «опасные указатели», соответствующий объект класса передаётся через шаблонные параметры (рисунок X)

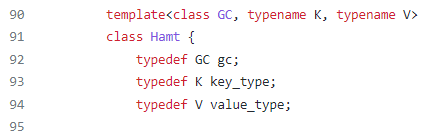


Рисунок 37 – Класс Hamt, принимающий в виде шаблона синглтон «опасный указатель»

В начале методов insert, lookup, remove создаётся фиксированный массив «опасных указателей (рисунок X)



Рисунок 38 - создание массива «опасных указателей»

Чтобы поменить указатель как «опасный», необходимо положить его в массив «опасных указателей» (рисунок X).

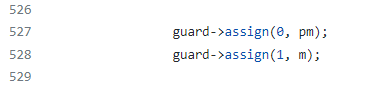


Рисунок 39 - установка «опасного указателя»

2.6 Внедрение в библиотеку libcds

Для внедрения в библиотеку libcds потребовалось изучение типовой структуры других контейнеров. Класс Hamt был расположен внутри namespace cds::container (рисунок X).

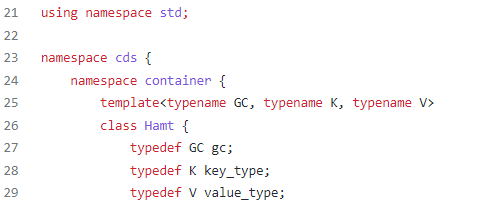


Рисунок 40 - расположение Hamt

Для использования встроенного класса Hamt необходимо произвести инициализацию инфраструктуры libcds (рисунок X).

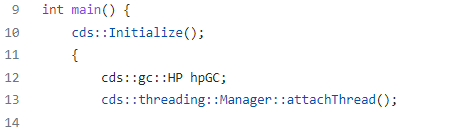


Рисунок 41 - инициализация инфраструктуры libcds

# 3 ТЕСТИРОВАНИЕ HASH ARRAY MAPPED TRIE

3.1 Модульное тестирование

Для внедрённого алгоритма были разработаны модульные тесты с помощью библиотеки googletest [26]. Часть из них проверяла однопоточную версию. В них были протестированы все сценарии использования методов классов Hamt и Bitmap.

Тесты были написаны по принципу arrange, act, assert (подготовить, действовать, проверить). Каждый тест делился на 3 части. В первой создаются объекты, необходимые в дальнейшем. Во второй вызывается метод, который нужно протестировать. В третьей, проверяются результаты после второго этапа. На рисунке X тестируется сокращение дерева. В начале создаётся объект Hamt. Далее добавляются три узла и удаляется тот, отсутствие которого должно вести к сокращению. Далее с помощь функции из библиотеки googletest проверяется структура дерева.

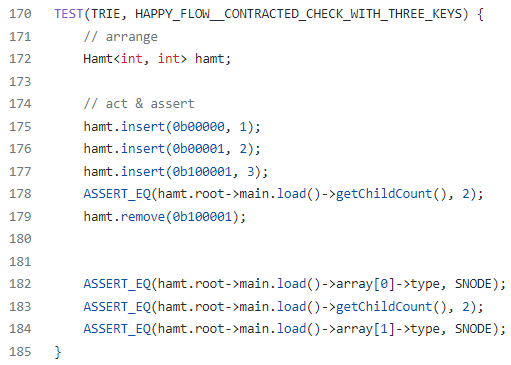


Рисунок 42 - Тестирование сокращения дерева

На рисунке X проверяется работа операции метода , который должен с помощью побитовых операций установить 0 на определённой позиции бита в числе.

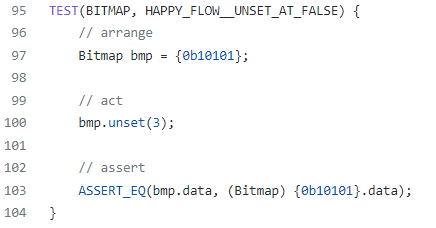


Рисунок 43 - Тестирование метода unset

Для тестирования многопоточных версий к предыдущим тестам были добавлены сценарии проверки на потокобезопасность. Базовые операции с деревом (insert, remove, lookup) запускались одновременно на множестве потоков. Несколько примеров сценариев:

* Вставка 100’000 значений одновременно 12 потоками
* Вставка 100’000 значений одновременно 12 потоками и последующее удаление всех значений
* Разное количество потоков выполняет одну из трёх операций (вставка, удаление, поиск) случайное количество раз.
* Предварительное заполнение структуры 100’000 ключей и одновременное удаление несколькими потоками.

Покрытие строчек кода было подсчитано при помощи утилиты gcov [27] и составило 94%.

3.2 Тестирование производительности

При помощи библиотеки google benchmark [28] были написаны тесты производительности. В качестве объектов тестирования были рассмотрены следующие структуры данных:

* однопоточная (потоконебезопасная) версия Hash Array Mapped Trie
* блокирующая многопоточная версия Hash Array Mapped Trie с грубой синхронизацией
* блокирующая многопоточная версия Hash Array Mapped Trie с тонкой синхронизацией
* многопоточная структура данных MichaelHashMap из библиотеки libcds
* set из стандартной библиотеки C++.

Для каждого теста воспроизводился один из сценариев:

* добавление множества ключей
* удаление из заполненного дерева множества ключей
* множественный поиск ключей в заполненном дереве

Операции производились 12 потоками, где у каждого была равная часть от общей работы. Для потоконебезопасных структур операции были окружены критическими секциями. Тестирование проводилось на машине со следующими характеристиками:

* 16 гигабайт оперативной памяти
* 64-разрядная система
* Intel® Core™ i7-9750H CPU @ 2.60GHz 2.59 GHz
* 12 логических процессоров
* 6 ядер

Для каждого тестового случая производилось 100 замеров скорости исполнения. Далее для результатов измерений были вычислены персентили (первый персентиль, второй персентиль, … сотый персентиль) и строился график зависимости времени исполнения (ось Y) от персентиля (ось X).

Разработанная структура данных lock-free Hash Array Mapped Trie сравнивалась с аналогичными структурами данных, которые могут поддерживать словарь ключ – значение. Это структуры множество (set) из стандартной библиотеки С++ и lock-free структура, основанная на хеш-таблице, MichaelHashMap. На рисунках 1, 2, 3 изображены персентили трёх структур для операции вставка, удаление, поиск. Для наглядности в рисунках X и X ось Y имеет логарифмическую размерность.

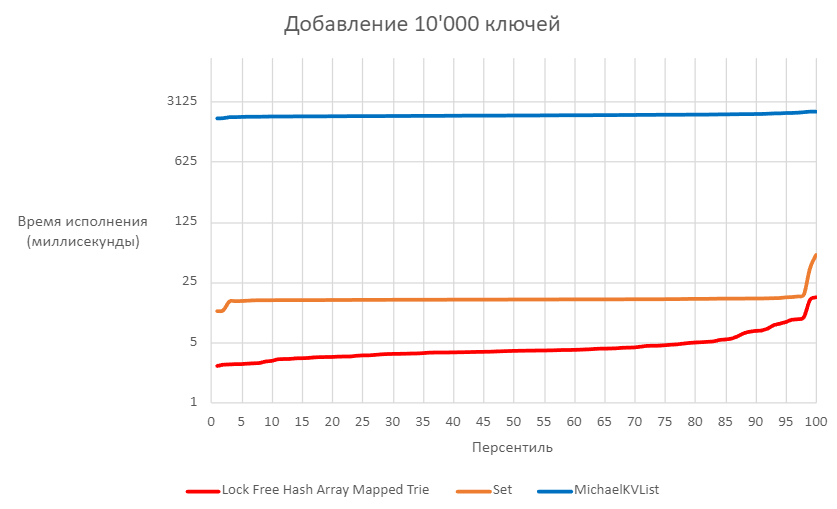


Рисунок 44 – График производительности при множественном добавлении

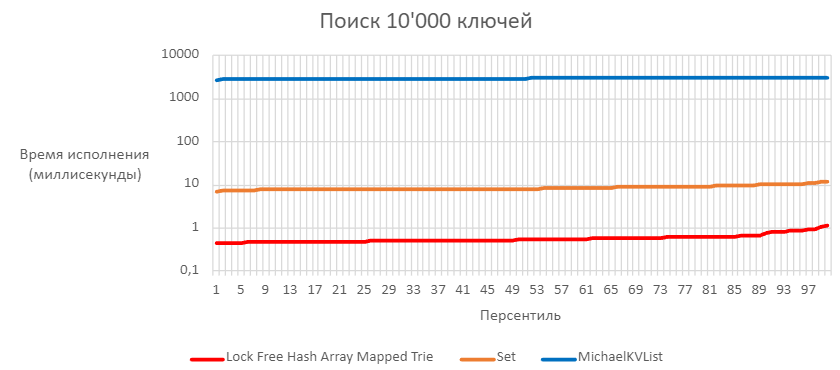


Рисунок 45 – График производительности при множественном поиске

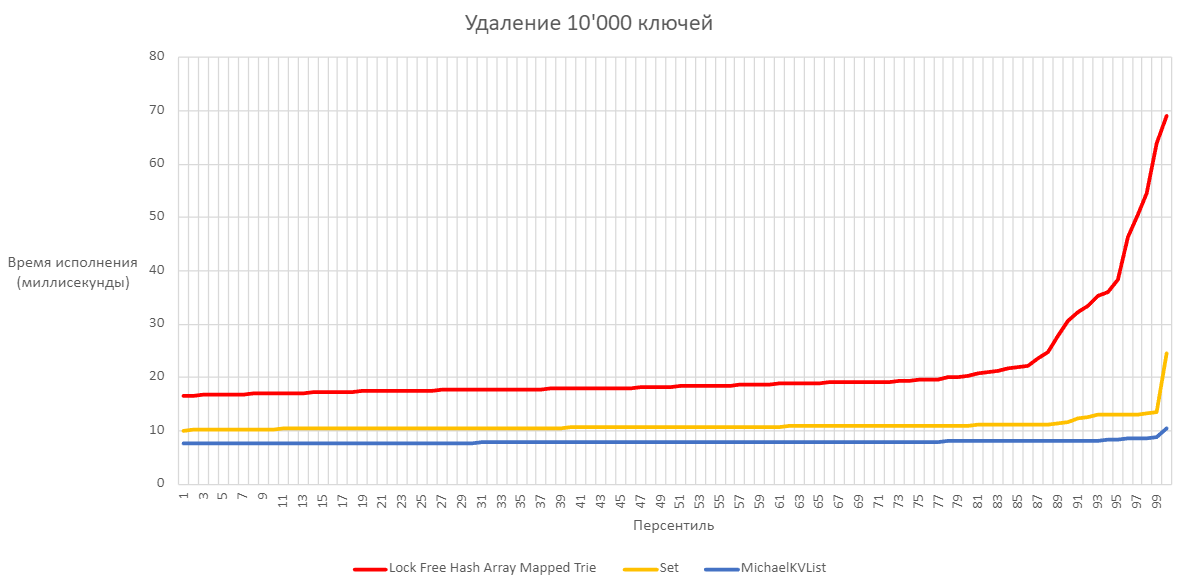


Рисунок 46 - График производительности при множественном удалении

График на рисунке X показывает, что с 95% вероятностью время исполнения структур данных

* lock-free Hash Array Mapped Trie будет меньше, чем 8.63 миллисекунд
* set будет меньше, чем 16.68 миллисекунд
* MichaelHashMap меньше, чем 2298.25 миллисекунд

Другими словами, скорость исполнения структуры данных lock-free Hash Array Mapped Trie будет на 51% быстрее set и на 99% быстрее MichaelHashMap.

Аналогично вычисляется, что при поиске lock-free Hash Array Mapped Trie быстрее set на 8 процентов, MichaelHashMap на 99%. При удалении Hash Array Mapped Trie показывает более долгое исполнение (на 73 % медленнее, чем Michael Hash Map).

Помимо сравнения с аналогами были проведены тесты lock-free Hash Array Mapped при разном количестве потоков (рисунок X) и сравнение с блокирующими и потоконебезопасной версиями (рисунок X).

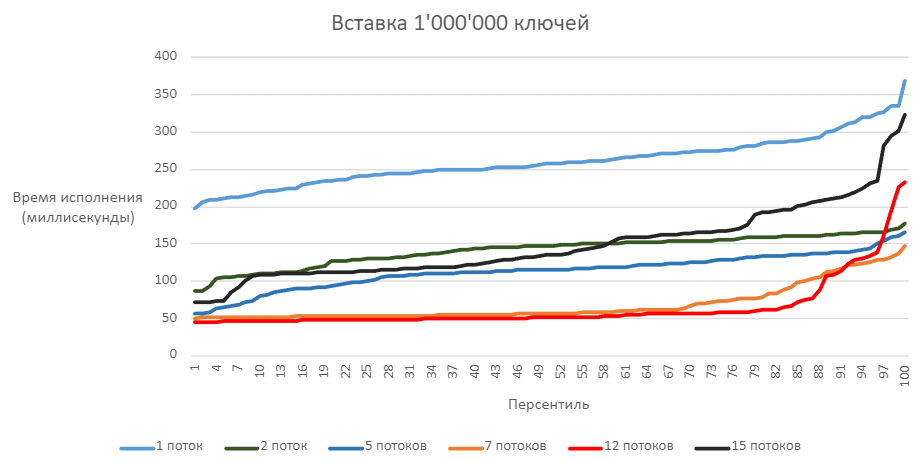


Рисунок 47 - График производительности при разном количестве потоков

Наиболее оптимальным количеством потоков для использования данной структуры данных является 12. На рисунке 42 видно, что с 90% вероятностью работа 12 потоков будет быстрее других конфигураций.

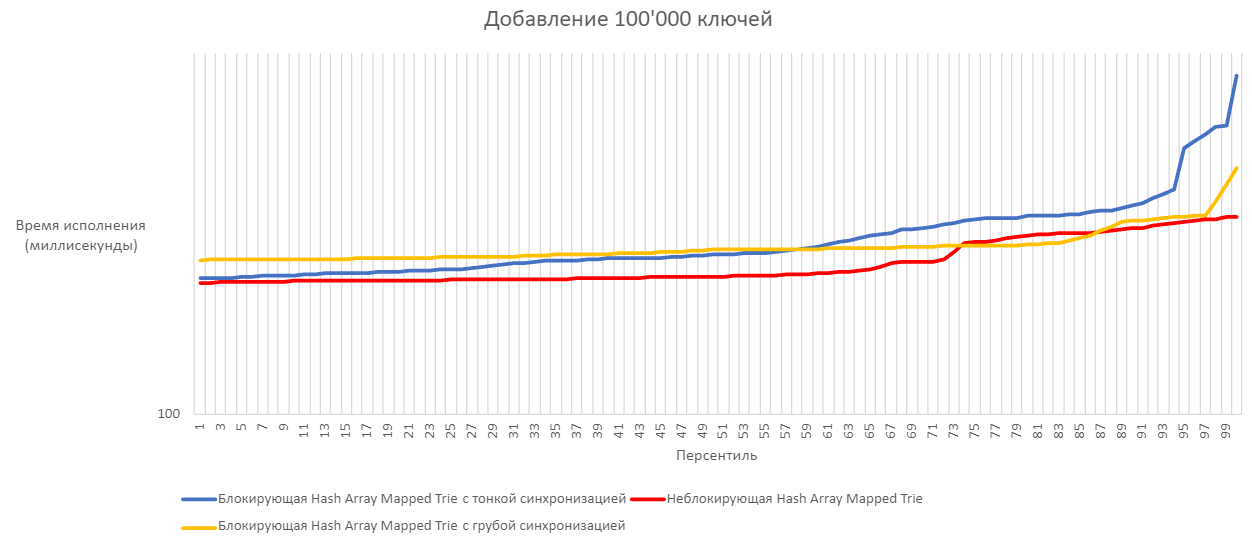
­­

Рисунок 48- График производительности разных версий HAMT

На рисунке X видно, что использование неблокирующей версии Hash Array Mapped Trie в 70% случаях быстрее блокирующих версий и потоконебезопасной.

# 4 ЗАКЛЮЧЕНИЕ

В данной работе удалось внедрить в библиотеку libcds lock-free структуру данных Hash Array Mapped Trie. Замеры производительности показали, что данная реализация при одновременном удалении множества ключей и при одновременном добавлении множества ключей показывает скорость исполнения ниже, чем в аналогичных структурах данных и однопоточной и блокирующей версии Hash Array Mapped Trie. Были протестированы основные сценарии исполнения и покрытие кода составило больше 90%.

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

x

|  |  |
| --- | --- |
| 1. | libcds [Электронный ресурс] // Github: [сайт]. [2015]. URL: https:/​/​github.com/​khizmax/​libcds (дата обращения: 15.4.2022). |
| 2. | Boost lock-free [Электронный ресурс] // boost: [сайт]. [2000]. URL: https:/​/​www.boost.org/​doc/​libs/​1\_60\_0/​doc/​html/​lockfree.html (дата обращения: 4.15.2022). |
| 3. | ConcurrencyKit [Электронный ресурс] // Github: [сайт]. [2000]. URL: https:/​/​github.com/​concurrencykit/​ck (дата обращения: 15.4.2022). |
| 4. | Штайн К., Ривест Р. Алгоритмы. Построение и анализ. 2019. 260-262 с. |
| 5. | DB-Engines Ranking of Key-value Stores [Электронный ресурс] // db-engines: [сайт]. URL: https:/​/​db-engines.com/​en/​ranking/​key-value+store (дата обращения: 15.4.2022). |
| 6. | Ахо А., Лам М., Сети Р., Ульман Д. Компиляторы. Принципы, технологии и инструментарий. 2016. 129-133 pp. |
| 7. | Хеш-таблица [Электронный ресурс] // neerc.ifmo.ru: [сайт]. [2003]. URL: https:/​/​neerc.ifmo.ru/​wiki/​index.php?title=%D0%A5%D0%B5%D1%88-%D1%82%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D0%B0 (дата обращения: 15.4.2022). |
| 8. | Prokopec A., Bagwell P., Odersky M. Cache Aware Lock-Free Concurrent Hash Tries 2011. URL: https:/​/​www.researchgate.net/​publication/​313432754\_Cache-Aware\_Lock-Free\_Concurrent\_Hash\_Tries (дата обращения: 4.15.2022). |
| 9. | Introduction to HAMT [Электронный ресурс] [2006]. URL: https:/​/​idea.popcount.org/​2012-07-25-introduction-to-hamt/ (дата обращения: 15.4.2022). |
| 10. | Bit array [Электронный ресурс] // Wikipedia: [сайт]. [2022]. URL: https:/​/​en.wikipedia.org/​wiki/​Bit\_array (дата обращения: 15.4.2022). |
| 11. | Hamming weight [Электронный ресурс] // Wikipedia: [сайт]. [2007]. URL: https:/​/​en.wikipedia.org/​wiki/​Hamming\_weight (дата обращения: 15.4.2022). |
| 12. | Atomic compare exchange [Электронный ресурс] // cppreference: [сайт]. [2009]. URL: https:/​/​en.cppreference.com/​w/​cpp/​atomic/​atomic\_compare\_exchange (дата обращения: 15.4.2022). |
| 13. | Git [Электронный ресурс] // git-scm: [сайт]. URL: https:/​/​git-scm.com/ (дата обращения: 15.4.2022). |
| 14. | Github [Электронный ресурс] // Github: [сайт]. [2000]. URL: https:/​/​github.com/ (дата обращения: 15.4.2022). |
| 15. | Actions [Электронный ресурс] // Github: [сайт]. [2018]. URL: https:/​/​github.com/​features/​actions (дата обращения: 15.4.2022). |
| 16. | Docker [Электронный ресурс] // Docker: [сайт]. [2008]. URL: https:/​/​www.docker.com/ (дата обращения: 15.4.2022). |
| 17. | Cmake [Электронный ресурс] // Cmake: [сайт]. [2007]. URL: https:/​/​cmake.org/ (дата обращения: 15.4.2022). |
| 18. | Make [Электронный ресурс] // gnu: [сайт]. [2009]. URL: https:/​/​www.gnu.org/​software/​make/ (дата обращения: 15.4.2022). |
| 19. | Clion [Электронный ресурс] // Jetbrains: [сайт]. [2000]. URL: https:/​/​www.jetbrains.com/​clion/ (дата обращения: 15.4.2022). |
| 20. | Iteration macros [Электронный ресурс] // Boost: [сайт]. [2000]. URL: https:/​/​www.boost.org/​doc/​libs/​1\_62\_0/​boost/​graph/​iteration\_macros.hpp (дата обращения: 4.15.2022). |
| 21. | Atomic [Электронный ресурс] // cppreference: [сайт]. [2000]. URL: https:/​/​en.cppreference.com/​w/​cpp/​atomic/​atomic (дата обращения: 15.4.2022). |
| 22. | Побитовые операции [Электронный ресурс] // tproger: [сайт]. [2019]. URL: https:/​/​tproger.ru/​translations/​bitwise-operations/ (дата обращения: 15.4.2022). |
| 23. | popcount [Электронный ресурс] // cppreference: [сайт]. [2000]. URL: https:/​/​en.cppreference.com/​w/​cpp/​numeric/​popcount (дата обращения: 4.15.2022). |
| 24. | ABA problem [Электронный ресурс] // Wikipedia: [сайт]. [2000]. URL: https:/​/​en.wikipedia.org/​wiki/​ABA\_problem (дата обращения: 15.4.2022). |
| 25. | Lock-free структуры данных. Внутри. Схемы управления памятью [Электронный ресурс] // Habr: [сайт]. [2017]. URL: https:/​/​habr.com/​ru/​post/​202190/ (дата обращения: 15.4.2022). |
| 26. | Googletest [Электронный ресурс] // Github: [сайт]. [2015]. URL: https:/​/​github.com/​google/​googletest (дата обращения: 15.4.2022). |
| 27. | Gcov [Электронный ресурс] // gcc: [сайт]. [2000]. URL: https:/​/​gcc.gnu.org/​onlinedocs/​gcc/​Gcov.html (дата обращения: 15.4.2022). |
| 28. | Google Benchmark [Электронный ресурс] // Github: [сайт]. [2009]. URL: https:/​/​github.com/​google/​benchmark (дата обращения: 15.4.2022). |

x