МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ Н.Э. БАУМАНА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

Метод блочного хранения данных с возможностью доказательства неправомерного доступа на основе хеш-сумм

Студент: Пересторонин Павел Геннадьевич

Группа: ИУ7-83Б

Руководитель: Григорьев Александр Сергеевич

Цель и задачи

Цель – разработать метод блочного хранения данных с возможностью доказательства неправомерного доступа с помощью хеш-сумм.

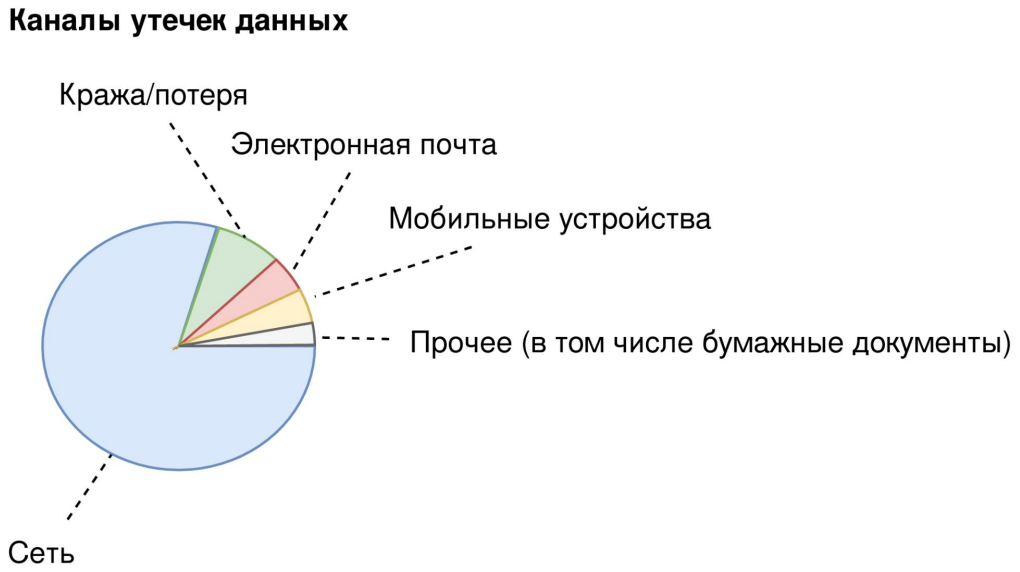
Задачи:

- рассмотреть базовые элементы и понятия, используемые при проектировании методов хранения информации с возможностью защиты от неправомерного доступа;
- провести анализ существующих методов хранения информации с защитой от неправомерного доступа;
- провести анализ блочного хранения данных на предмет защиты информации от неправомерного доступа;
- спроектировать и реализовать метод блочного хранения данных с возмозможностью доказательства неправомерного доступа;
- исследовать метод на предмет невозможности реализации угроз при различных конфигурациях системы.

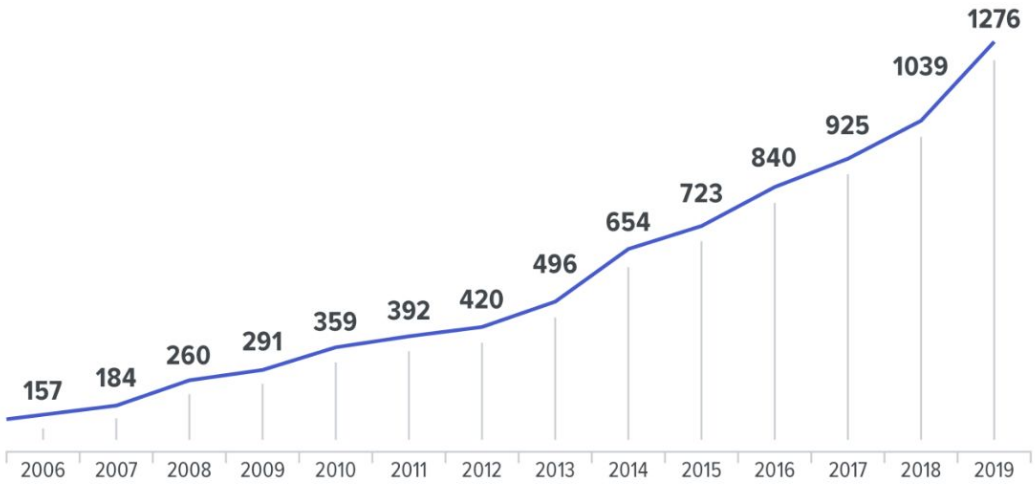
Защита информации ограниченного доступа

Информация ограниченного доступа:

- коммерческая тайна;
- персональные данные;
- служебная тайна;
- секрет производства;
- другое.



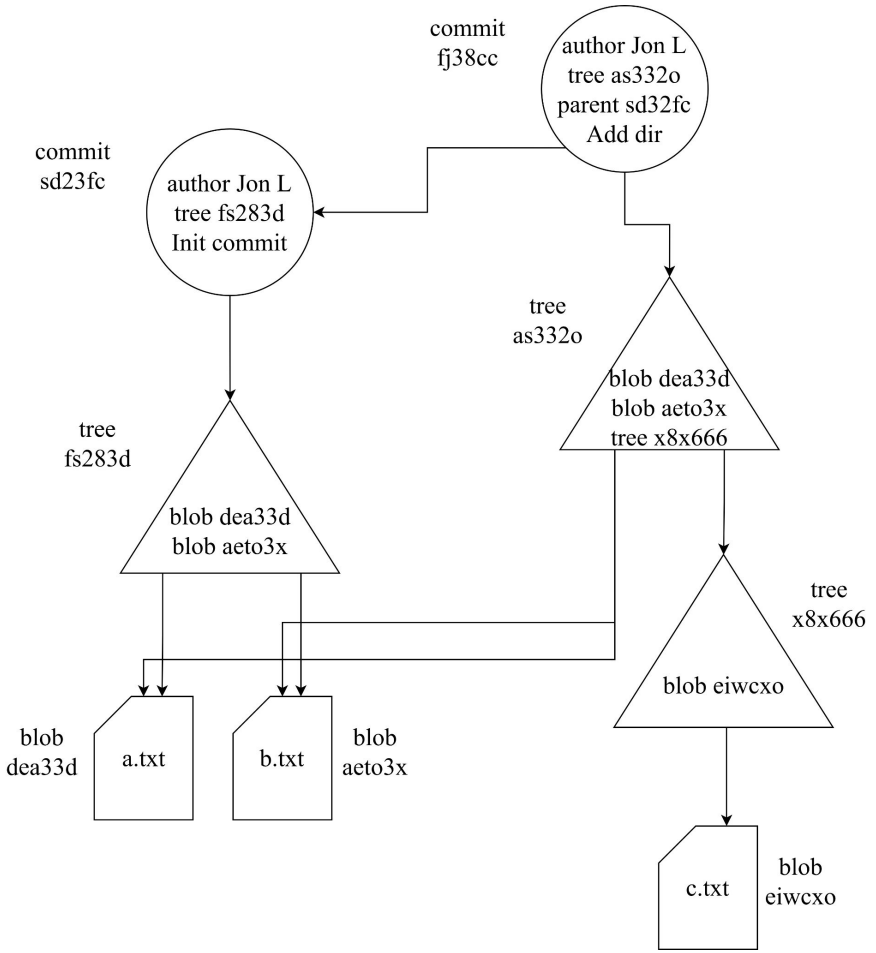
- количество утечек увеличивается;
- большую долю составляет сеть;
- тенденция на ужесточение регулирования в сфере информационной безопасности со стороны государств в мире.



Методы локального хранения данных с защитой от неправомерного доступа

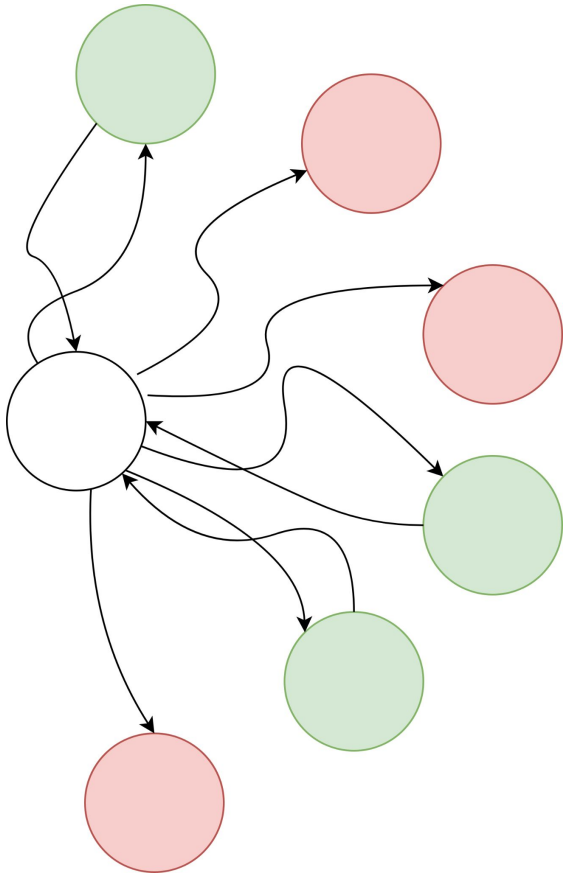
Криптографические файловые системы

Git

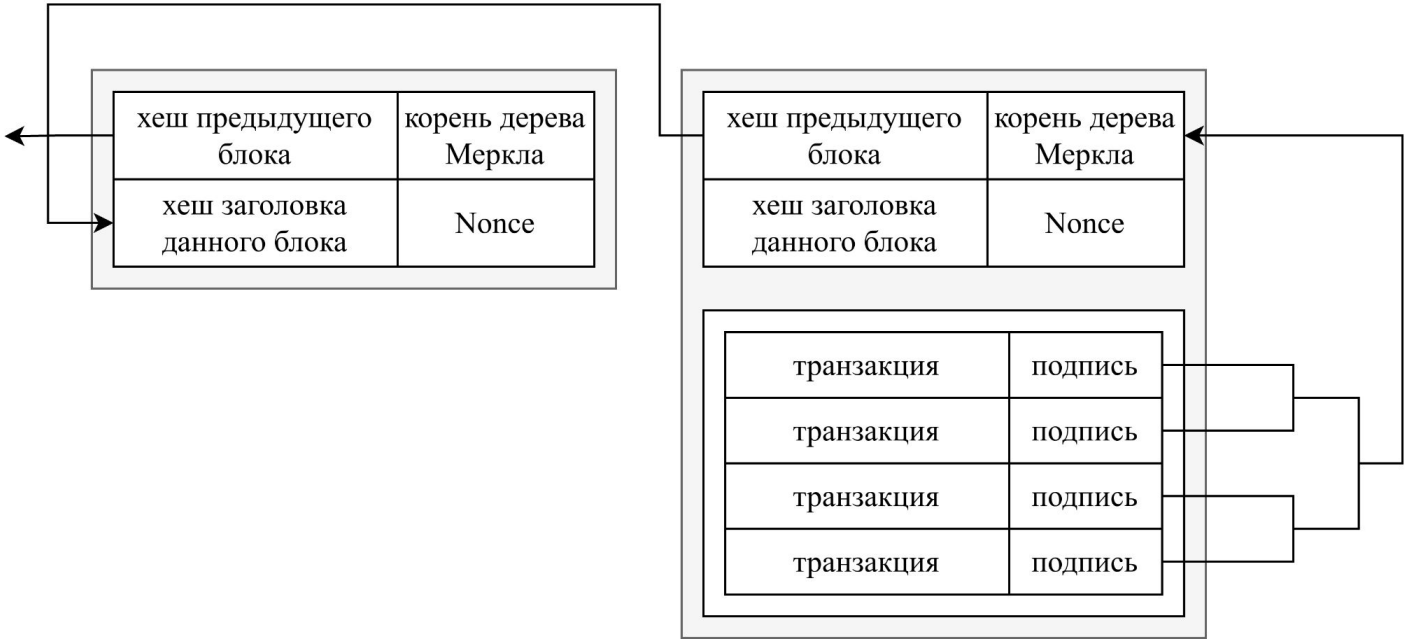


Методы распределенного хранения данных с защитой от неправомерного доступа

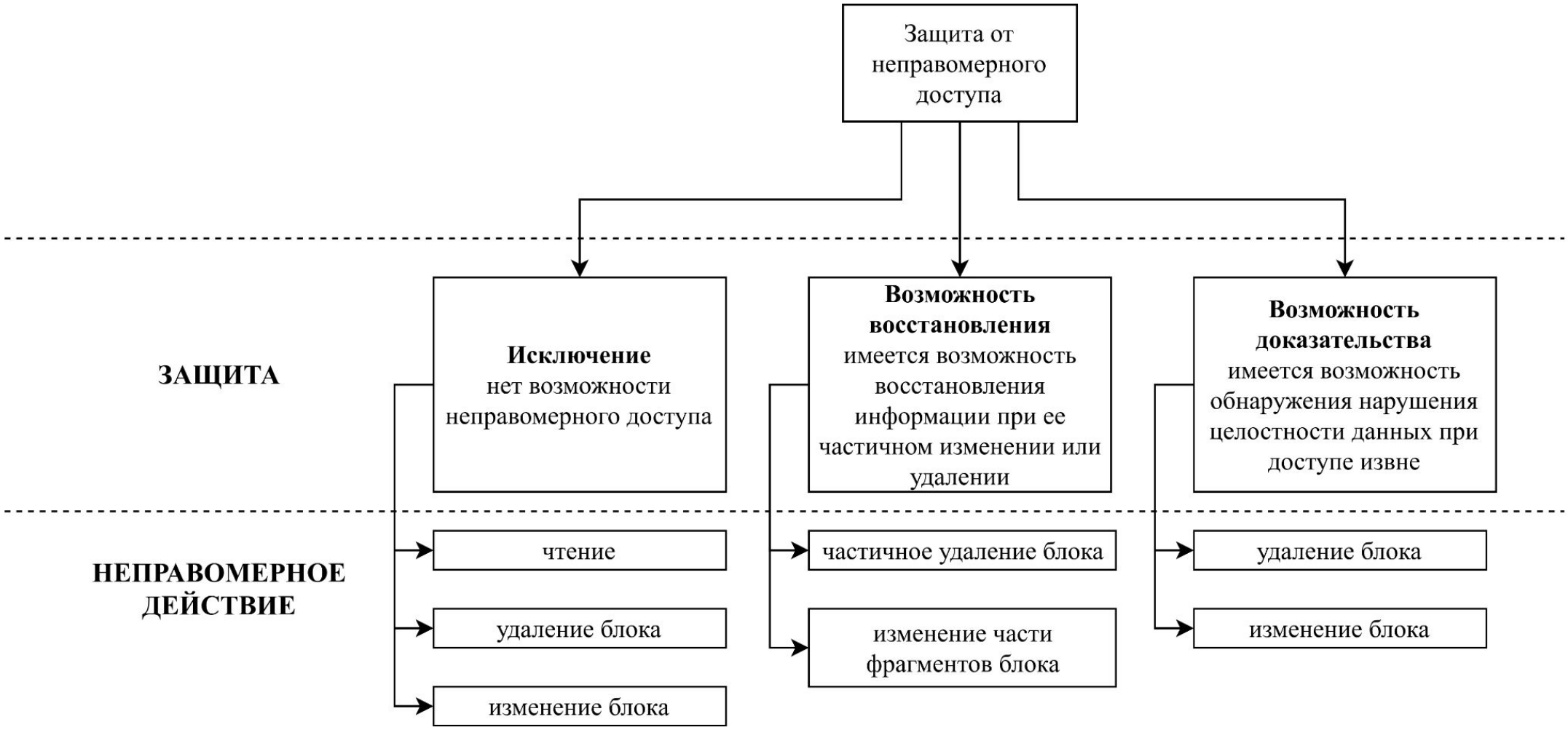
PASIS и OceanStore



Bitcoin



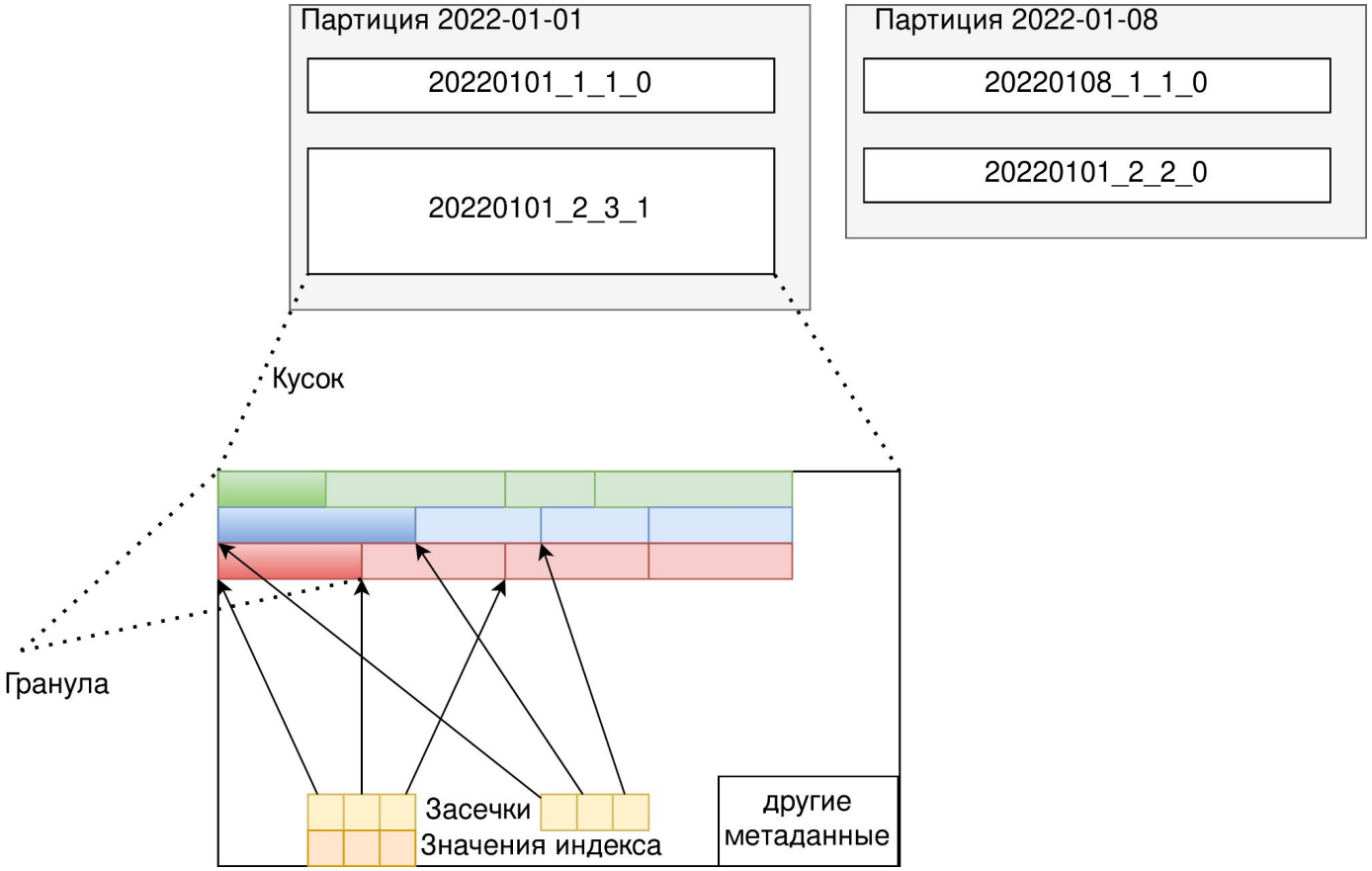
Виды защиты от неправомерного доступа



Анализ существующих решений на предмет защиты от неправомерного доступа

неправомерное действие (защита)	**PASIS**	**КФС**	**OceanStore**	**Git**	**Bitcoin**
чтение (исключение)	-	+	-	-	-
частичное изменение/удаление (восстановление)	+	-	+	-	-
изменение/удаление (исключение)	-	-	-	-	+
изменение (доказательство)	+	-	+	+/-	+
удаление (доказательство)	-	-	-	+/-	+

Метод блочного хранения данных в СУБД



СУБД с блочным хранением данных:

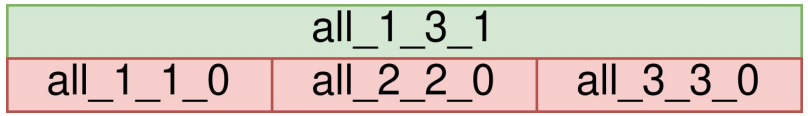
- Oracle Exadata;
- Vertica;
- **ClickHouse**.

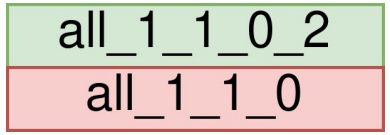
Компоненты СУБД ClickHouse:

- Звено — блок, единица хранения информации.
- Партиция — логическая группа звеньев.
- Гранула — единица записи и чтения данных.
- Индекс — отсортированные значения первых в гранулах первичных ключей.
- Засечки — смещение столбцов в файле для значений индекса.

Операции с данными в СУБД ClickHouse в MergeTree

- Вставка:
  - каждая вставка — новый блок;
  - атомарность операции за счет использования временного блока и переименовывания.
- Слияния:
  - служат для оптимизации хранения и поиска;
  - на входе произвольное количество кусков;
  - на выходе всегда 1 кусок.
- Мутации:
  - служат для изменения данных;
  - могут применяться по несколько к одному куску;
  - в одной задаче мутации участвует 1 кусок.



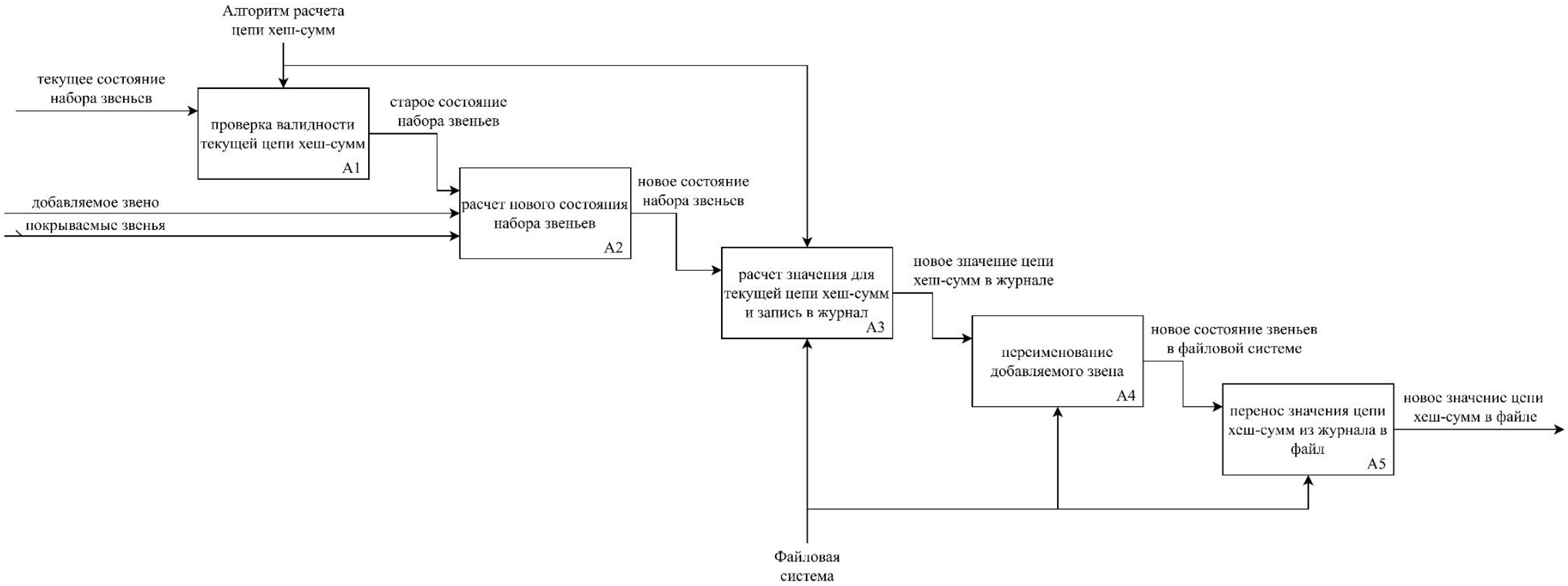


Анализ существующей защиты данных от неправомерного доступа в движке MergeTree

- Дополнительные возможности:
  - шифрование данных на уровне директории и столбца;
  - проверка целостности данных на уровне звена.

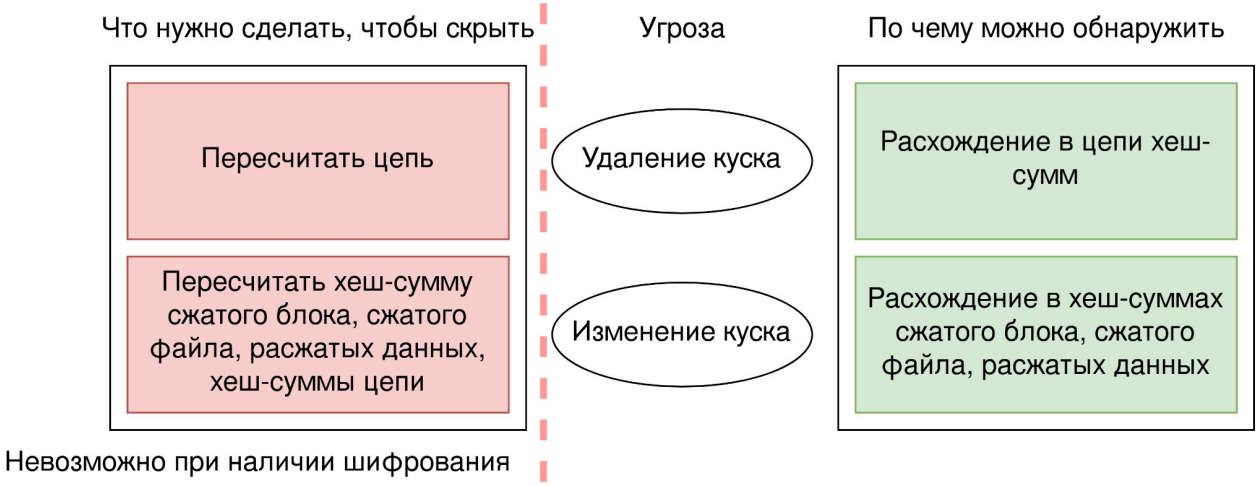
неправомерное действие (защита)	**без шифрования**	**с шифрованием**
чтение (исключение)	-	+
удаление/изменение (исключение)	-	-
частичное удаление/изменение (восстановление)	-	-
удаление блока (доказательство)	-	-
изменение (доказательство)	+/-	+

Функциональная модель программного комплекса



Предлагаемый метод хранения данных с возможностью доказательства неправомерного доступа

Сценарии возможных угроз:



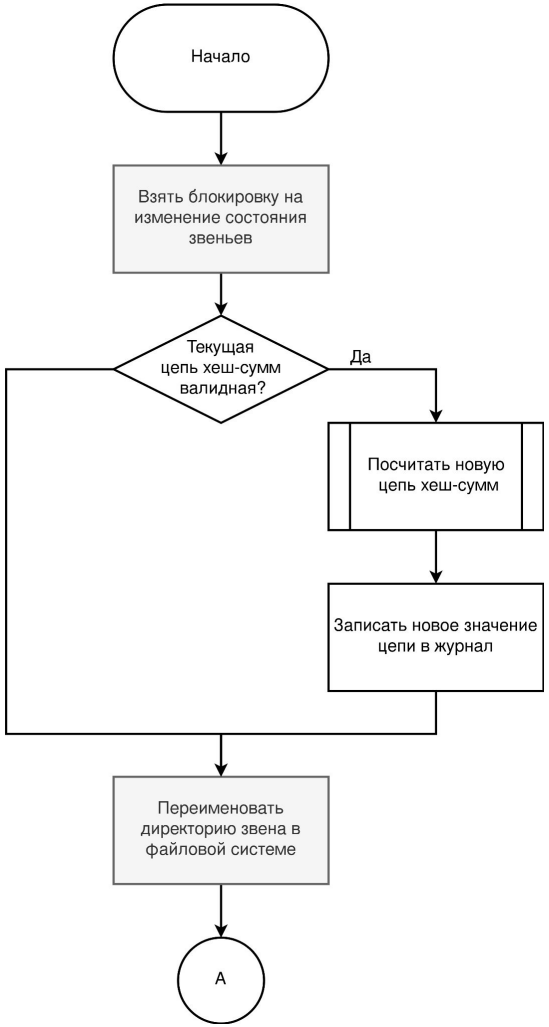
Формула расчета цепи хеш-сумм:

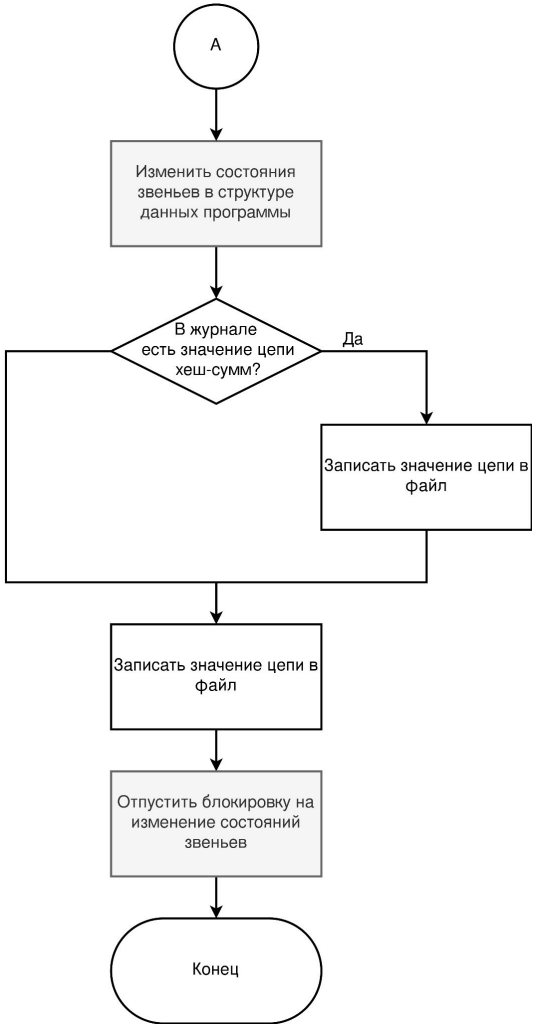
$$y_i = hash(x_i|y_{i-1}), y_1 = hash(x_1)$$

где:

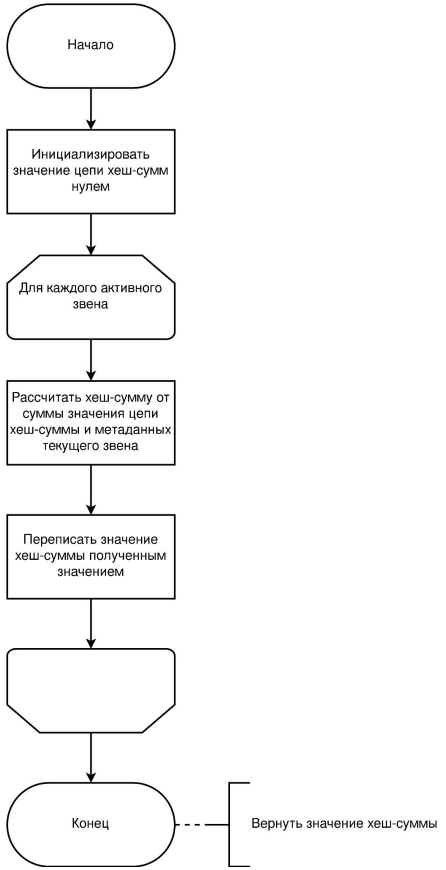
- i — индекс звена;
- hash — хеш-функция;
- | — операция конкатенации байтовых массивов;
- x — метаданные звена в байтовом представлении.

Методы расчета и валидации цепи хеш-сумм и обновления состояния звеньев

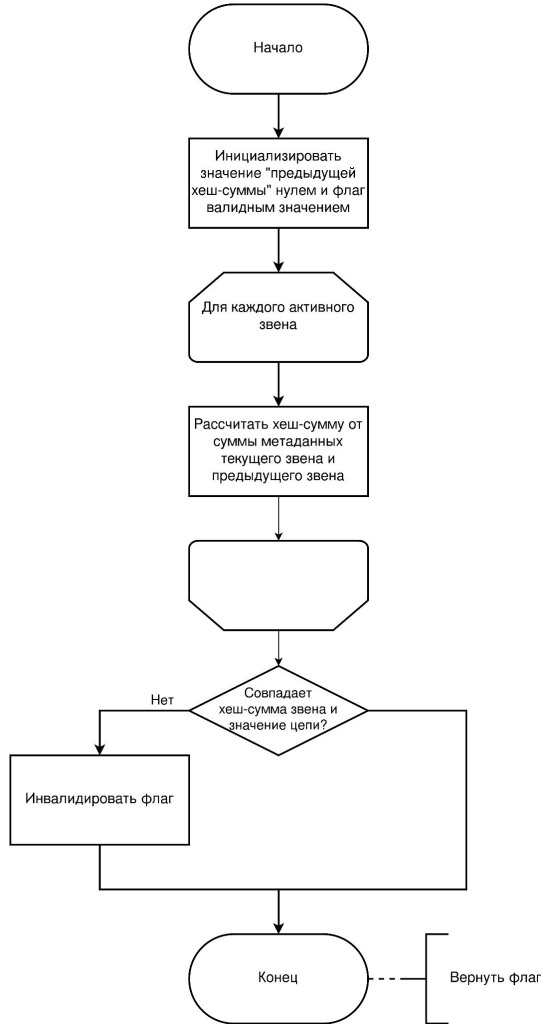




обновление состояния звеньев



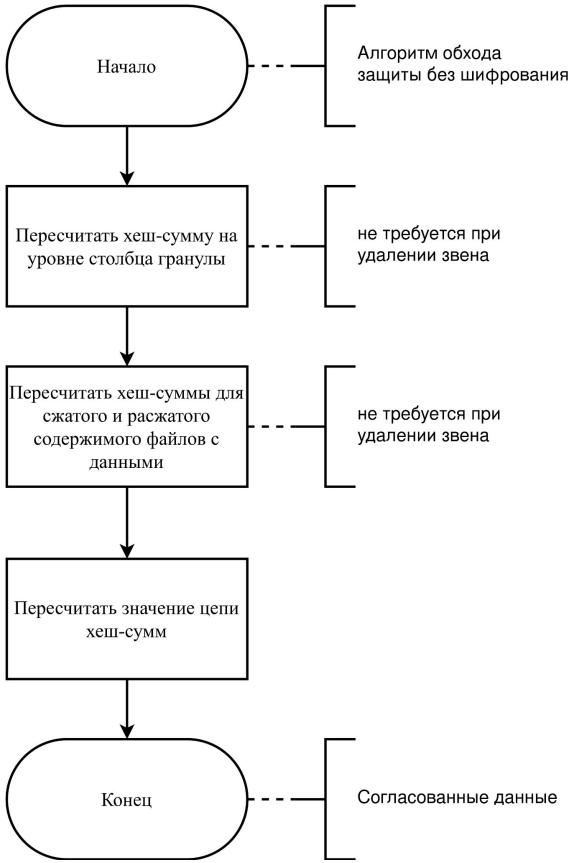
расчет цепи хеш-сумм



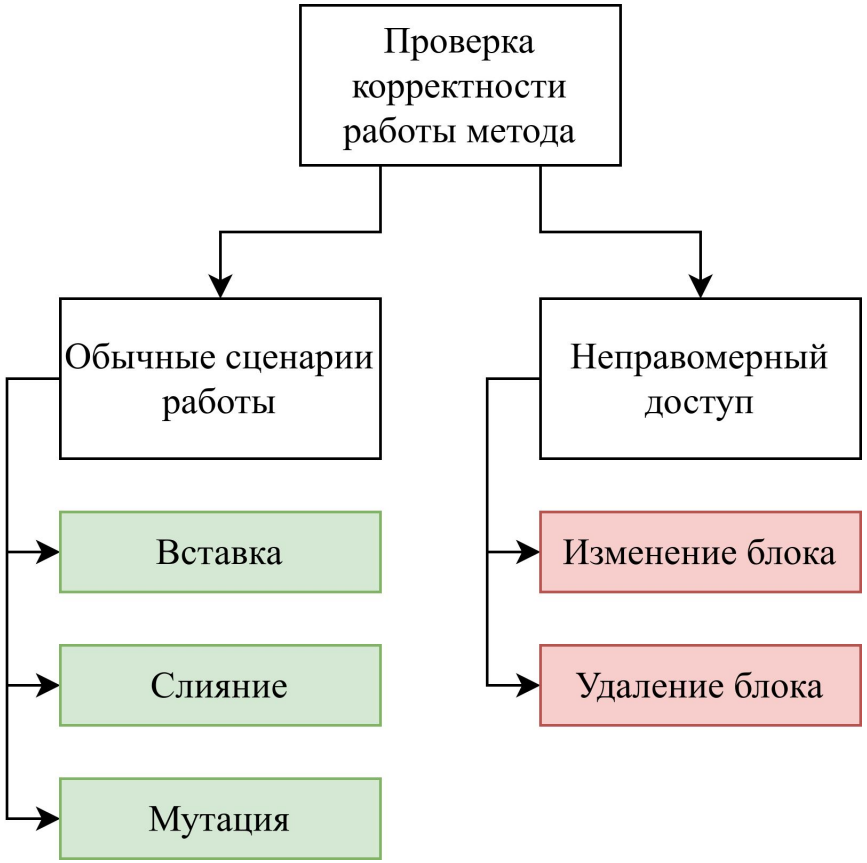
валидация цепи хеш-сумм

Исследование метода на предмет невозможности реализации угроз при различных конфигурациях системы

Без шифрования:



С шифрованием:



Заключение

В результате выполнения данной работы была достигнута цель работы, а также решены все поставленные задачи, а именно:

- были рассмотрены базовые элементы и понятия, используемые при проектировании методов хранения информации с возможностью защиты от неправомерного доступа;
- был проведен анализ существующих методов хранения информации с защитой от неправомерного доступа;
- был проведен анализ блочного хранения данных на предмет защиты информации от неправомерного доступа;
- был спроектирован и реализован метод блочного хранения данных с возмозможностью доказательства неправомерного доступа;
- метод был исследован на предмет невозможности реализации угроз при различных конфигурациях системы.

Направления дальнейшего развития

- реализация метода блочного хранения данных с возможностью доказательства неправомерного доступа для движка ReplicatedMergeTree СУБД ClickHouse;
- реализация возможности восстановления после частичного удаления или изменения звена в движке ReplicatedMergeTree СУБД ClickHouse.

ReplicatedMergeTree — аналогичный с точки зрения физического хранения данных MergeTree движок, обладающий возможностью репликации звеньев.