Министерство образования и исследований Республики Молдова

Бельцкий государственный университет им. Алеку Руссо

Факультет реальных наук, экономики и окружающей среды

Кафедра математики и информатики

**Способы хранения информации в базах данных**

Научный реферат

Выполнил:

Михай Валентин, гр. АW11M

Бэлць, 2021

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 3](#__RefHeading___Toc659_347429684)

[ОСНОВНЫЕ ПОНЯТИЯ 4](#__RefHeading___Toc661_347429684)

[КЛАССИФИКАЦИЯ СПОСОБОВ ХРАНЕНИЯ ИНФОРМАЦИИ В БАЗЕ ДАННЫХ 4](#__RefHeading___Toc663_347429684)

[ЗАКЛЮЧЕНИЕ 6](#__RefHeading___Toc665_347429684)

[БИБЛИОГРАФИЯ 7](#__RefHeading___Toc667_347429684)

# ВВЕДЕНИЕ

Бурное развитие информационных технологий и информационных систем всё чаще отражается во всё большем количестве сфер деятельности человека. Программные продукты предоставляют автоматизированные и удобные решения для задач, которые ранее требовали ручного труда или усилий группы работников.

На ряду с ростом популярности цифровых решений и развитием технологий в общем, также развиваются технологии, которые позволяют улучшать отдельные части программы. Одним из классов таких технологий являются инструменты для хранения и обработки данных.

На сегодняшний день практически каждое приложение (программное обеспечение) использует одну из реализаций баз данных для обработки и хранения информации. Количество хранимой информации в банках различных продуктов растет экспоненциально, так, например, компания Google в 2006 году использовала базу на 850 терабайт для своей поисковой машины, а уже в 2014 году объем информации вырос до 10 эксабайт [1]. Также каждая задача, поставленная перед программистом, затрагивает различные доменные области, а соответственно информацию различных типов и форматов.

Для работы с различными типами данных и разнообразием форматов базы данных стали развиваться в узконаправленных направлениях: для хранения информации о финансовых операциях следует использовать безопасную базу с поддержкой транзакций, в то время как для хранения кэша приложения лучше подойдет решение, которое использует оперативную память компьютера.

В зависимости от реализации и типа хранения информации базы данных можно классифицировать по нескольким критериям:

* модель данных;
* среда постоянного хранения;
* тип содержимого;
* степень распределенности.

Данные параметры позволяют характеризовать большинство случаев и реализаций методов хранения информации в базе. Однако, существуют и такие задачи, которые не поддаются классификации, например сверхбольшие базы данных.

Такие базы данных представляют собой наборы информации критических объемов, которые определены научными и техническими достижениями в сфере физического хранения информации.

# ОСНОВНЫЕ ПОНЯТИЯ

База данных, согласно стандарту ISO/IEC 2382:2015, - это совокупность данных, организованных в соответствии с концептуальной структурой, описывающей характеристики этих данных и взаимоотношения между ними, которая поддерживает одну или более областей применения [2].

Система управления базами данных(СУБД) – это набор программ, которые обеспечивают управление созданием и использованием баз данных [3].

# КЛАССИФИКАЦИЯ СПОСОБОВ ХРАНЕНИЯ ИНФОРМАЦИИ В БАЗЕ ДАННЫХ

В общем случае, способы хранения информации можно характеризовать четырьмя основными критериями: модель данных, средства постоянного хранения, тип содержимого, степень распределенности.

**Модели данных**

По типам моделей данных способы хранения могут быть следующими:

* иерархическая;
* объектно-ориентированная;
* реляционная.

Иерархическая модель представляет данные в виде древа, которое состоит из отдельных объектов, между которыми существуют связи.

Объектно-ориентированная модель является отражением принципов объектно-ориентированного программирования и хранит информацию в виде классов, их атрибутов и методов [4].

Реляционная модель – это логическая модель, где данные хранятся исходя из их связей и отношений между отдельными сегментами.

**Средства постоянного хранения**

Для хранения информации может быть использован любой тип памяти компьютера:

* вторичная память – это энергонезависимая, периферийная память, например, жесткий диск;
* оперативная память – используется резидентными базами данных;
* третичная память – представляет собой съемное устройство независимое от конкретного сервера базы данных.

**Типы содержимого**

В зависимости от типа хранимой информации относительно предметной области способы организации данных также отличаются. Основными группами типов можно считать:

* документы – информация организована в виде документа и содержит исчерпывающее описание;
* мультимедиа – вся информация представлена медиа содержимым;
* клиентская – различные виды личной информации различных типов данных и правилами контроля.

**Степени распределенности**

База данных не обязательно должна быть централизованной. Ее составные части могут быть распределены по разным серверам. Выделяют следующие типы:

* централизованная – для поддержки базы данных используется один компьютер;
* Распределенная – поддержка базы делегирована нескольким машинам:
  + однородная – каждая машина хранит схожие сегменты данных и использует одинаковые СУБД
  + неоднородная – для поддержки используется более одной СУБД, соответственно сегменты данных могут различаться по типу.

В качестве отдельных способов хранения информации можно выделить: репликация и шардинг.

Репликация позволяет хранить несколько копий текущего состояния базы балансировать нагрузку внутри кластера приложения [5]. Существуют две основные модели репликации: мастер – мастер, мастер – подчененный. В первом случае все реплики являются равносильными, во втором – подчиненные реплики имеют более низкий приоритет.

Шардинг представляет собой подход. при котором информация разбивается на несколько частей на уровне одной сущности [6]. В случае базы данных, которая использует табличную структуру, одна таблица может быть разбита на две более мелкие, что повысит скорость обработки каждого отдельного сегмента.

# ЗАКЛЮЧЕНИЕ

Способ хранения информации данных в базе напрямую связан с характеристиками и метаданными самих этих данных. На сегодняшний день существует большое количество различных способов хранения, которые позволяют обрабатывать самые разные типы данных, от географических координат до реквизитов банковских карт.

Некоторые подходы позволяют повысить отказоустойчивость базы, например: репликация. Другие позволяют манипулировать большими объемами информации, используя шардинг. Такие методы позволяют повысит безопасность системы и ее отзывчивость.

Также для каждой специфической реализации потребителя данных существуют различные модели, которые повышают быстродействие системы и делают ее использование более удобным.

Несмотря на всё изобилие способов хранения информации существуют ограничения физических аппаратов, которые приводят к вопросу о проблеме сверхбольших баз данных. На сегодняшний день данная проблема не решена полностью, в то время как скорость накопления данных растет очень быстро.

# БИБЛИОГРАФИЯ

1. CHITU, ALEX. *How Much Data Does Google Store?* [online] [цитирован 29.09.2021] Доступен: http://googlesystem.blogspot.com/2006/09/how-much-data-does-google-store.html
2. *ISO/IEC 2382:2015 Information technology — Vocabulary* [online] [цитирован 29.09.2021] Доступен: https://www.iso.org/standard/63598.html
3. *ISO/IEC TR 10032:2003 Information technology — Reference Model of Data Management* [online] [цитирован 29.09.2021] Доступен: <https://www.iso.org/standard/38607.html>
4. КУЗНЕЦОВ, СЕРГЕЙ*. Объектно-ориентированные базы данных - основные концепции, организация и управление: краткий обзор* [online] [цитирован 29.09.2021] Доступен:<http://citforum.ru/database/articles/art_24.shtml>
5. CROCETTI, PAUL. *Data replication* [online] [цитирован 29.09.2021] Доступен:<https://searchdisasterrecovery.techtarget.com/definition/data-replication>
6. ЗОРИН, ОЛЕГ. *Шардирование баз данных* [online] [цитирован 29.09.2021] Доступен: http://oleg.zorin.ru/doku.php?id=development:database:sharding