

OLAP СИСТЕМА АНАЛИЗА

Никита МАКЕЕВ

Департамент Программная Инженерия и Автоматика, группа TI-196, Факультет Вычислительной Техники, Информатики и Микроэлектроники, Технический Университет Молдовы, Кишинев, Молдова

Автор Корреспондент: Никита МАКЕЕВ, e-mail: macheev.nikita@isa.utm.md

Научный руководитель: Дориан САРАНЧУК, DISA, FCIM, UTM

Аннотация: Статья посвящена OLAP системе анализа обработки данных. Описана общая информация о системе OLAP. Типы OLAP систем, пример с кубом OLAP, различные операции над кубом и разобрано понятие тест FASMI.

Ключевые слова: СУБД, OLAP, Тест FASMI, Типы OLAP.

Введение

В нашем мире все больше и больше фирм, предприятий, корпораций работают с большими объемами данных, имеют своего рода приложения, которые нужны для того, чтобы комплексно проанализировать данные и провести различного рода манипуляции. Данный анализ необходим из-за того, что в итоге он должен поспособствовать принятию различных решений. Для этого служащие должны обладать данной информацией и зачастую, количественной информацией. Именно поэтому такой набор данных следует собрать из всех источниках систем предприятия, а после чего проанализировать. Именно для этого делаются хранилища данных и в свою очередь система OLAP помогает анализировать и обрабатывать большие массивы данных.

OLAP определение

OLAP – это интерактивная аналитическая обработка, своего рода технология обработки данных, основная суть которой заключается в подготовке общей информации на основе больших данных, которые структурированы по многомерному принципу [1].

Основоположник термина OLAP — Эдгар Кодд, в основе его концепции лежит куб. OLAP куб – это структура данных, оптимизированная для очень быстрого анализа данных.

Тест FASMI

С течением времени определение Кодда переросло в тест FASMI, в свою очередь который требует, чтобы OLAP-приложение представляло своеобразную возможность быстро проанализировать информацию.

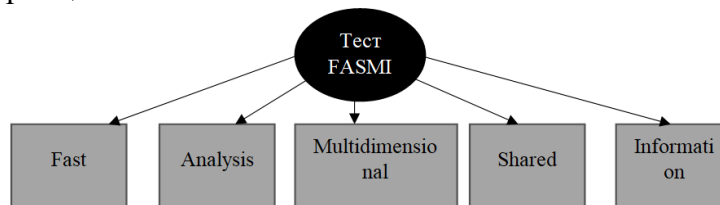


Рисунок 1. Тест FASMI.

Fast (Быстрый) – анализ должен происходить быстро по всем источникам информации. Примерно - 5 с или менее.

Analysis (Анализ) – должна быть возможность осуществлять своего рода анализ например, числовой и статистический.

Shared (Разделяемой) – большое количество пользователей может иметь доступ к данным, при этом необходимо контролировать доступ к конфиденциальной информации.

Multidimensional (Многомерной) – любая информация должна быть многомерной. Это основная, наиболее существенная характеристика OLAP [2].

Information (Информации) – приложение должно иметь возможность обращаться к любой нужной информации, независимо от ее объема и места хранения.

Типы OLAP

Существует 3 типа OLAP

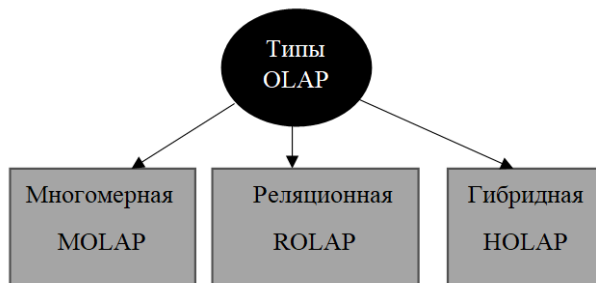


Рисунок 2. Типы OLAP.

MOLAP – Одна из форм OLAP. Основная суть которой заключается в суммировании базы данных и создании многомерной схемы данных с возможностью сохранения, как базовых данных, так и агрегатов [3].

ROLAP – также один из типов OLAP, который работает с реляционной базой данных по итогу все измерения, факты, агрегаты хранятся в реляционных таблицах.

HOLAP – тип OLAP, который использует реляционные таблицы для хранения базовых данных и многомерные таблицы для агрегатов.

Куб OLAP

Куб OLAP – это структура данных с гарантированной возможностью обеспечения быстрого анализа данных. Кубы могут отображать и производить различные операции с объемными данными, в свою очередь предоставляя пользователям доступ к данным и возможность поиска, благодаря чему все данные могут быть сведены в фрагменты и обработаны по мере нужды для поиска решений в различных вопросах [4].

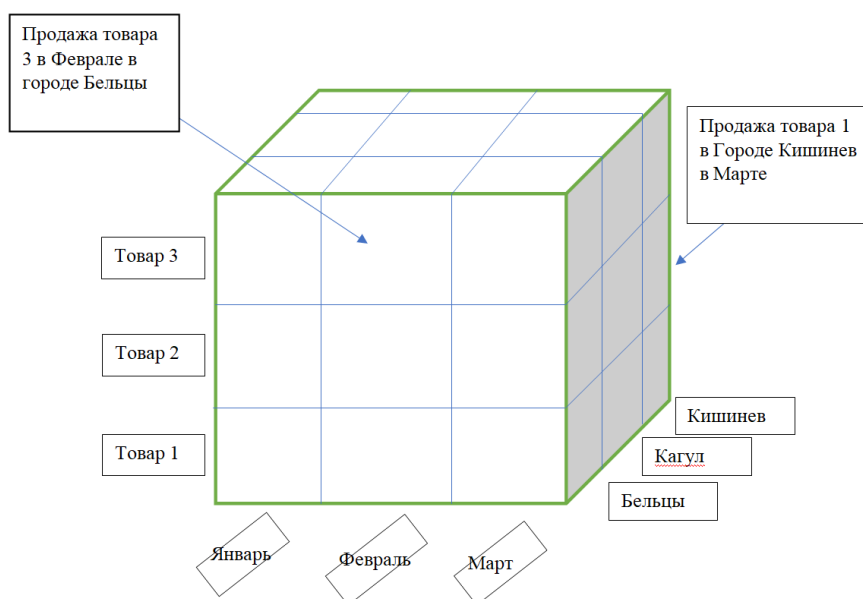


Рисунок 3. Пример Гиперкуба для продаж в Кишиневе

Операции над Гиперкубом

1) Свернуть

Свернуть – это когда данные, которые разделены, объединяются, а затем используются в дальнейшем. Это можно сделать благодаря:

- Уменьшению размеров.
- Используя иерархию понятий, где система группирования вещей делается на основе определенного порядка или уровня.

Пример

Предположим, что наш магазин развивается и выходит на рынок СНГ

Таблица 1

Пример Свертывания

Прибыль за Январь	Кишинев
Прибыль за Январь	Кагул
Прибыль за Январь	Бельцы
Прибыль за Март	Киев
Прибыль за Март	Одесса

Таблица 2

Таблица после свертывания

Прибыль за Январь	Молдова
Прибыль за Март	Украина

2) Развертывание

Развертывание – это своего рода разделение данных на мелкие части, которое также применяется к кубу данных. В примере указанном в таблице 3 применяется расширение измерений, расширение это не что иное, как добавление новых измерений к тому что имеем. Также есть условие, что если данные существуют то мы можем их расширить. Также расширение можно применять к любой оси куба данных [5].

Процесс может быть сделан:

- Переход вниз по ожидаемой иерархии при фрагментации.
- Увеличение размерности текущих наборов данных.

Пример в Таблице 2 мы получили прибыль за месяц в определенном магазине в Кишиневею

Таблица 3

Таблица после Развертывания

Прибыль товара 1	Магазин Linella
Прибыль товара 1	Магазин Nr1
Прибыль товара 1	Магазин Metro

3) Разрез

Разрез используется для того, чтобы получить двумерное представление куба

Так как у нас размерность куба больше 3 мы можем получить различное множество срезов [6]

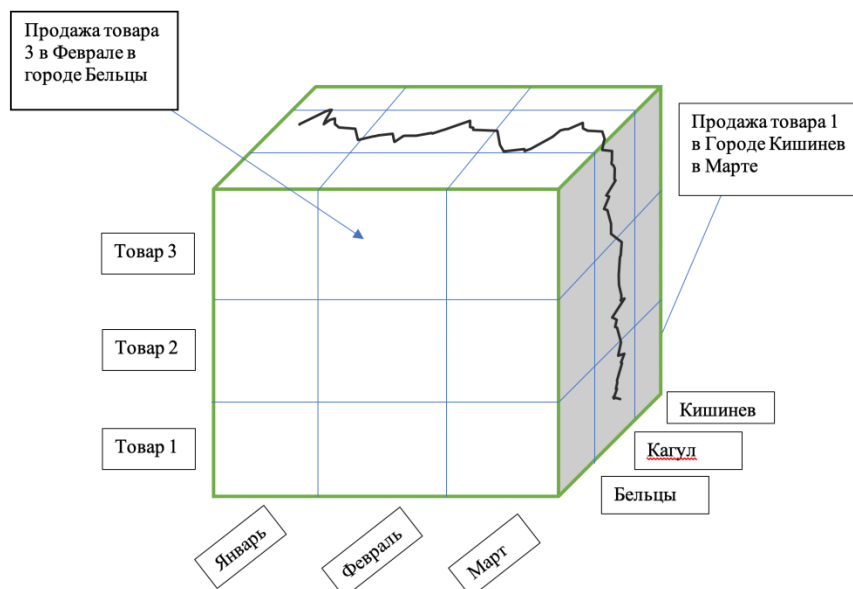


Рисунок 4. Двумерный срез куба для одной меры

Таблица 4

Таблица после Развертывания

	Январь	Февраль	Март
Товар 1	10	3	5
Товар 2	4	5	7
Товар 3	6	5	4

В приведенном примере мы сделали двумерный срез куба для одной меры а именно ‘Кагул’ и получили таблицу 4, где видно сколько и какого товара было продано в каждый месяц [7]

4) Вращение

Вращение – это своего рода ориентации куба различных данных для проверки других представлений, которые могут иметь данные. С помощью поворачивания одного измерения куба данных, мы получаем реализацию операции вращения. Так для нашего примера мы изменили область товара с оси X на ось Y и месяц с оси Y на ось X, таким образом, предоставив ей другое представление.

Таблица 5

Таблица после Развертывания

	Товар 1	Товар 2	Товар 3
Январь	10	4	6
Февраль	3	5	5
Март	5	7	4

Заключение

В информационном мире, который с каждым годом все больше и больше развивается и процесс информатизации набирает колоссальный рост, всю информацию нужно хранить в различных базах данных. Миру просто необходимо решение в виде OLAP системы, у которой в свою очередь очень значимые плюсы такие как, предметная ориентированность, прямой доступ к данным, удобная навигация по данным, визуализация данных, неизменность данных, Высокая скорость формирования отчетов и др. И это лишь часть из большого количества преимуществ использования OLAP системы, все это в совокупности дает нам отличную систему благодаря которой, можно очень быстро проанализировать различные данные и принять оперативные решения.

Библиография

1. Microsoft Docs. Куб OLAP- [online]. [20.12.2022]. – Доступно: <https://docs.microsoft.com/ru-ru/system-center/scsm/olap-cubes-overview?view=sc-sm-2019> .
2. Крмс . Структура OLAP Системы- [online]. [20.12.2022]. – Доступно: https://www.krms.ru/Automatization/OLAP_system.htm .
3. Олар . Введение в OLAP и многомерных баз данных- [online]. [20.12.2022]. – Доступно: <http://www.olap.ru/basic/alpero2i.asp> .
4. Coderlessons . Введение в OLAP и многомерных баз данных- [online]. [20.12.2022]. – Доступно: <https://coderlessons.com/tutorials/bolshie-dannye-i-analitika/teoriia-khraneniia-dannykh/7-chto-takoe-olap> .
5. Wikipedia. OLAP - [online]. [20.12.2022]. – Доступно: <https://ru.wikipedia.org/wiki/OLAP>.
6. Habrhabr. Введение в многомерный анализ. - [online]. [20.12.2022]. – Доступно: <https://habr.com/ru/post/126810/>.
7. Education-wiki. Операции в OLAP. - [online]. [20.12.2022]. – Доступно: <https://ru.education-wiki.com/1424365-operations-in-olap> .