# СИСТЕМА АГРЕГИРОВАНИЯ ДАННЫХ ДЛЯ ФИНАНСОВОЙ АНАЛИЛТИКИ

## Миронов Д. С.

Научный руководитель: д.т.н., проф., зав. БК №234 УЭВМ, Парамонов Николай Борисович

#### РТУ МИРЭА, Институт информационных технологий

**Аннотация:** В статье автор описывает построение системы позволяющую быстро внедрять аналитические методы и принимать решения на основе данных

**Ключевые слова:** OLAP, куб, аналитика, индекс, данные.

Современные финансовые системы генерируют огромное количество данных, требующих оперативной обработки и представления в удобной для анализа форме. В данной статье описан подход к построению системы агрегирования данных на основе OLAP-кубов, ориентированной на задачи финансовой аналитики. Предложена архитектура и методы, позволяющие эффективно агрегировать, индексировать и визуализировать многомерные финансовые данные.

Финансовая аналитика требует не только точности, но и высокой скорости получения агрегированных показателей. Принятие решений на основе данных становится основным стандартом для организаций. Однако традиционные ВІ-инструменты, такие как Qlik Sense или Hyperion Planning, зачастую не соответствуют требованиям гибкости, скорости внедрения и стоимости владения.

Создание собственной системы агрегирования данных, ориентированной на финансовые показатели, позволяет достичь высокой адаптивности под задачи конкретного предприятия или аналитической группы. В основе предлагаемой системы лежит технология OLAP (On-Line Analytical Processing), позволяющая формировать многомерные кубы для быстрого анализа больших объемов информации.

В данной статье описывается построение системы позволяющую быстро внедрять аналитические методы и принимать решения на основе данных. В виду того что в создаваемой системе основной функционал будет построен на создании OLAP – куба.

#### Пример функционального решения.

Перед началом формирование OLAP - куба, необходимо создать его структуру (рис. 1), то из чего он будет состоять. Основой, конечно же являются данные и стороны. Стороны — это измерение куба, то что будет группироваться, при формировании куба.

Необходимо указать столбцы основного файла и их иерархию. Пример иерархии или же одной стороны — «месяц — неделя — год». Данная сторона будет называться в структуре, например, «Дата». Таких сторон в кубе может быть не ограниченное количество, но с каждой добавляемой стороной и глубины иерархии, увеличивается объем куба и сложность при его расчете. В качестве данных необходимо указать столбцы с числовыми значениями на основе которых буду проводиться расчеты.

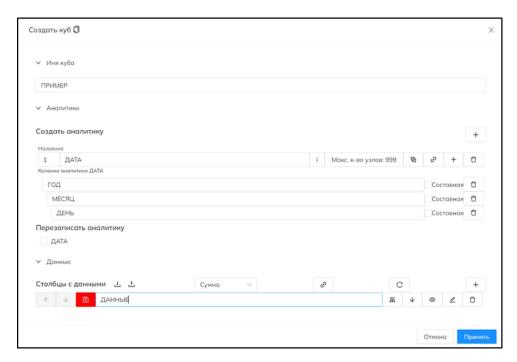


Рис. 1. Формирование структуры [разработано автором]

После описания сторон и данных куба, формируются параметры каждой стороны и столбов с данными, для того чтобы эффективно хранить полученные значения и быстро выводить данные при запросе.

Для каждой стороны в структуре должна находится следующая информация: название стороны, названия столбцов иерархии стороны, порядковый номер стороны, длина индекса для данной стороны.

Требуемая память для сборки куба составляет (1):

$$x * 3 = y \tag{1}$$

 $\Gamma$ де x — размер загруженного массива данных, у — требуемая память системы.

## Получение разворота OLAP-куба по индексам.

Для того чтобы получить запрашиваемый разворот, необходимо сделать запрос, в котором указан индекс запрашиваемой стороны, если запрашиваемый разворот включает в себя две стороны, то в запросе нужно указать два индекса.

Приведем пример (рис. 2), для каждой стороны присвоен индекс, а максимальная длина каждого два знака, количество символов зависит от количества уникальных значений в столбце стороны, «ID чел» - 1, «ДАТА» - 2, «Данные» - 0 и 1, столбец с данными в индексе всегда стоит последним и должен иметь минимум два значения индекса, для суммы всех столбцов с данными (это индекс 0) и для каждого столбца индекс становится плюс 1. Так как в примере всего лишь один столбец с данными, индекс будет выглядеть следующем образом 010100, на рисунке 3, представлен разворот с индексами 000100 и 010100.

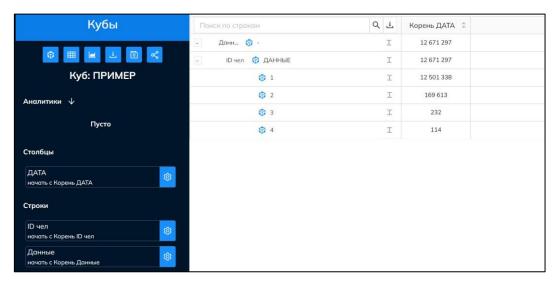


Рис. 2. Разворот куба [разработано автором]

Так же может понадобиться рассмотреть столбец «ДАТА» более детально, то нажав на соответственный столбец, появятся уровни данной стороны (рис. 3).



Рис. 3. Разворот куба с детальным разворотом столба [разработано автором]

Так как все данные рассчитаны на этапе формирования куба, скорость получения данных составляет O(1).

## Литература:

- 1. Введение в OLAP и многомерные базы данных URL: http://www.olap.ru/basic/alpero2i.asp
- 2. «I Was Seduced By a Build Scenario»: 11 Ways to Avoid This Exec's Greatest Tech Failure//Better Cloud // https://www.bettercloud.com/monitor/build-vs-buy/