[](http://www.google.by/url?sa=i&rct=j&q=&esrc=s&frm=1&source=images&cd=&cad=rja&docid=bGt2gk_80A6EGM&tbnid=UQw5JkCd-NLRZM:&ved=0CAUQjRw&url=http%3A%2F%2Fwww.orchidltd.co.uk%2F&ei=bk3xUfKEKcm6OPZT&bvm=bv.49784469,d.ZWU&psig=AFQjCNGGXXJtTG8_0SI5hSRL7ZllnezSWA&ust=1374854887632211)

**MiBroKi: Money Transfer**

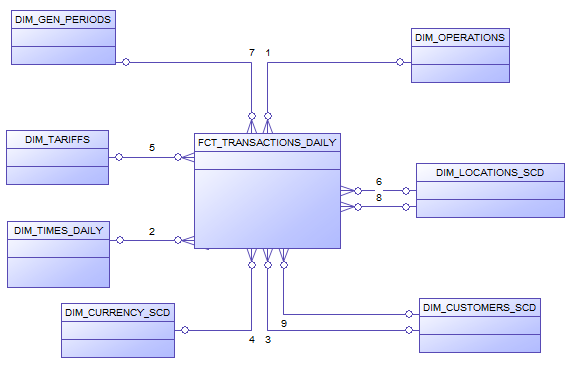
**BI Solution Proposal**

**VERSION NUMBER 1.0**

|  |  |
| --- | --- |
| **Submission Date:** | *7/26/2013* |
| **Requested By:** | *Kyril Bucha* |
| **Business Owner:** | *Irina Brodetskaya* |
| **Contact Info:** | *Iryna\_Bradzetskaya@epam.com* |

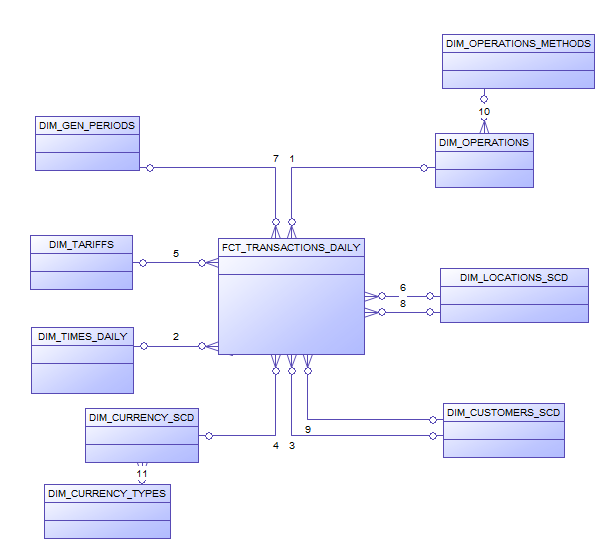
# DWH Solution Concept

## Star Logical Diagram

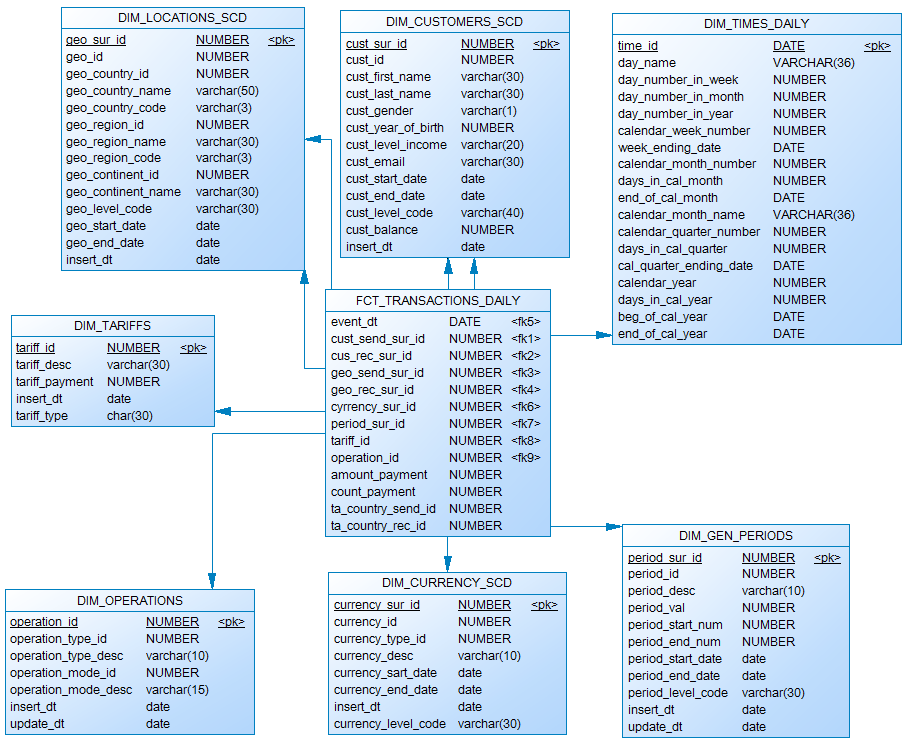


|  |  |
| --- | --- |
| N | LOGICAL SCHEMA |
| 1 | FCT\_TRANSACTIONS\_DAILY is a fact table. It’s get information from all other dimensions and transaction characteristics |
| 2 | DIM\_CUSTOMERS\_SCD – there is full information about user |
| 3 | DIM\_TIEMS\_DAILY – List of all possible periods for analysis (Year, month and quarters) |
| 4 | DIM\_LOCATIONS\_SCD - list of all possible divisions by countries |
| 5 | DIM\_OPERATIONS - full information about operations types |
| 6 | DIM\_CURRENCY\_SCD full information about of all possible currency name and types |
| 7 | DIM\_TARIFFS – full information about tariffs and types of them |
| 8 | DIM\_GEN\_PERIODS - contains information about the various periods and intervals used in the system. |

## Logical SNOW Flakes schema



## Star Physical diagram



## Dimensions Types Description

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N | Name | Type | Size | DW – Merged Dimensions | Descriptions |
| 1 | **DIM\_TIMES\_DAILY** | SCD1 | BIG | DW.T\_DAYS, DW.T\_WEEKS, DW.T\_MONTHS, DW.T\_QUARTERS,  DW.T\_YEARS | Dimension describe time parameters of the business |
| 2 | **DIM\_LOCATIONS\_SCD** | SCD2 | SMALL | DW.T\_COUNTRIES  DW.T\_CNTR\_GROUPS  DW.T\_CNTR\_SUB\_GROUPS  DW.LC\_CNTR\_GROUPS  DW.T\_GEO\_TYPES  DW.T\_GEO\_SYSTEMS  DW.LC\_GEO\_SYSTEMS  DW.T\_GEO\_PARTS  DW.T\_GEO\_REGIONS  DW.T\_GEO\_OBJECTS  DW.T\_CNTR\_GROUP\_SYSTEMS  DW. LC\_CNTR\_GROUP\_SYSTEMS  DW.LC\_CNTR\_SUB\_GROUPS  DW.LC\_ GEO\_PARTS  DW.LC\_COUNTRIES  DW.LC\_ GEO\_REGIONS | Dimension table with full information about all countries, regions, subregions and some other classifications. |
| 3 | **DIM\_CUSTOMERS\_SCD** | SCD2 | BIG | DW.T\_CUSTOMERS  DW.T\_GENDER  DW.T\_LEVEL\_INCOME  DW.T\_EMAIL | Dimension contains detailed information about each user of system. |
| 4 | **DIM\_CURRENCY\_SCD** | SCD2 | SMALL | DW.T\_CURRENCY  DW.T\_CURRENCY\_TYPES | Type dimension with list of all currencies and types of them. |
| 5 | **DIM\_GEN\_PERIODS** | SCD1 | SMALL | DW.T\_PERIOD\_DESC  DW.T\_PERIOD\_START  DW.T\_PERIOD\_END  DW.T\_LEVEL\_CODE | Dimension contains information about the various periods and intervals used in the system. |
| 6 | **DIM\_TARIFFS** | SCD1 | SMALL | DW.T\_TARIFFS  DW.T\_TARIFFS\_TYPE | This dimension contains list of all tariffs of company and types of them. |
| 7 | **DIM\_OPERATIONS** | SCD1 | SMALL | DW.T\_OPERATIONS\_TYPES  DW.T\_OPERATIONS\_METHODS | Dimension table contains details information about operations and used methods for them. |

## Dimensions Hierarchies

**DIM\_TIMES\_DAILY:**

**Hierarchy DAY-WEEK-MONTH-YEAR**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| DAY | DAY | Store day at the calendar | DAY\_ID |
| WEEK | WEEK | Store weeks at the calendar year | WEEK\_ID |
| MONTH | MONTH | Store months at the calendar year | MONTH\_ID |
| YEAR | YEAR | Store years at the calendar year | YEAR\_ID |

**Hierarchy DAY--MONTH- QUARTER -YEAR**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| DAY | DAY | Store day at the calendar year | DAY\_ID |
| MONTHS | MONTH | Store months at the calendar year | WEEK\_ID |
| QUARTER | QUARTER | Store quarters at the calendar year | QUARTER\_ID |
| YEAR | YEAR | Store years at the calendar year | YEAR\_ID |

**DIM\_CURRENCY\_SCD:**

**Hierarchy CURRENCY –- CURRENCY TYPE**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| CURRENCY | CURRENCY | Stores information about currencies | CURRENCY \_ID |
| CURRENCY TYPES | CURRENCY\_TYPE | Stores information about different types of currencies | CURRENCY\_TYPE\_ID |

**DIM\_LOCATIONS\_SCD:**

**Hierarchy COUNTRY –SUBREGION-REGION-CONTINENT**

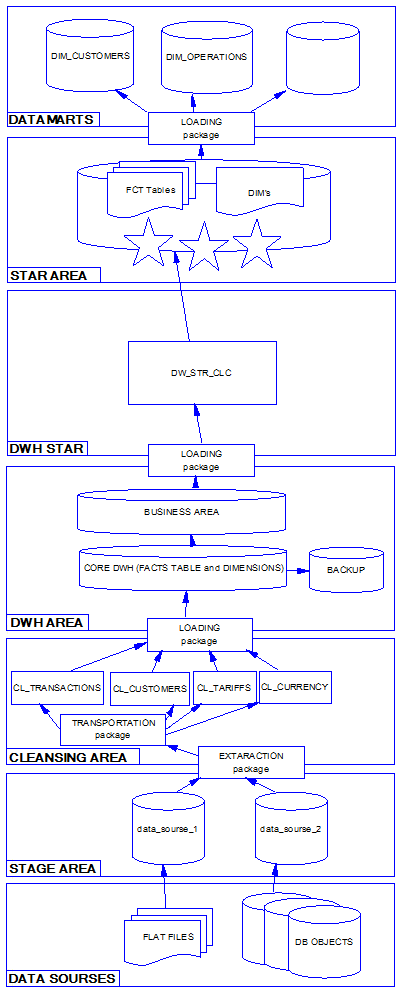
|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| COUNTRIES | GEO\_COUNTRY\_NAME | Store countries for each region. | GEO\_COUNTRY\_ID |
| REGIONS | GEO\_REGION\_NAME | Store regions of the world. | GEO\_REGION\_ID |
| CONTINENT | GEO\_CONTINENT\_NAME | Store continents of the world. | GEO\_CONTINENT\_ID |

## Facts Aggregations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N | Name | Code | Table Name | Additive | Descriptions |
| 1 | Number of transactions | FCT\_COUNT\_TRNSCT | FCT\_ COUNT\_TRNSCT | + | Calculate distinct values of Transaction at the EVENT\_DT period. |
| 2 | Amount of money by operations | FCT\_AMOUNT | FCT\_AMOUNT | + | Calculate amount of money conducted by each type of operations at the EVENT\_DT period. |

## DWH architecture

|  |  |  |  |
| --- | --- | --- | --- |
| Level Type | Object Name | Tablespace | Description |
| Storage level | SA\_CUSTOMERS  **User\_name:** U\_SA\_CUSTOMERS | ts\_sa\_customers\_data | Contains information about customers, transactions, and information necessary for the correct operation of the system. Users of this level should be able to load data from source files, create tables and view in this area. Upload data from source files to these tables and views and send it to next area. |
| SA\_TRANSACTIONS  **User\_name:** U\_SA\_TRANSACTIONS | ts\_sa\_trnsct\_data |
| SA\_DATA  **User\_name:** U\_SA\_DATA | ts\_sa\_data |
| DW - Cleansing Level | DW\_CL  **User\_name:**  U\_DW\_CL | ts\_dw\_cl\_data | U\_DW\_CL clean raw data inserted from files and insert it into tables in the 3-rd normal form. To make this operations user U\_DW\_CL should have grants to create tables with raw data, to connect and select information from resources and load information, and create views, to create result views with clean information. |
| DW – Level | DW  User\_name: U\_DW | ts\_dw\_data | Store information in the 3-rd normal form. |
| DW– Prepare Star Cleansing Level | SAL\_DW\_CL  **User\_name:** U\_SAL\_DW\_CL | ts\_sal\_dw\_cl\_data | Create views, which consist of merged tables from DW Level. U\_SAL\_CL\_DW should have grants to connect and select information from DW Level, create tables, where he will make changes, and create any view to make the clean view visible to next level users, who will have grants on it. |
| STAR - Cleansing | SAL\_CL  User\_name: U\_SAL\_CL | ts\_sal\_cl\_data | Delete from views, from previous level information, which don’t touch our business analysis question. There should be only data with are useful for our analysis. U\_SAL\_CL should have grants to connect and select information from the view, which was created on U\_SAL\_CL\_DW level, create tables to insert into them the information, which relative only with future analytic plan and the grant to create views to make public clean information for future levels. |
| STAR – Level | FCT\_TRANSACTIONS  User\_name: U\_FCT\_TRANSACTIONS | ts\_star\_fct\_data | Receiving data from STAR – Cleansing Area. Store information about facts. User should be able to select data from tables and view from previous level and create tables and view for data marts. |
| DIM\_TABLES  User\_name: U\_DIM\_TABLES | ts\_star\_dim\_data | Receiving data from STAR – Cleansing Area. Store information about dimensions. User should be able to select data from tables and view from previous level and create tables and view for data marts. |

* 1. **Dataflow Diagram**
  2. **Partitioning**

The query execution speed, which works with fact table, can be increased with partitioning.

As business plan of the star is to show quickly the information about transactions by operations and in section of time, so the composite partitioning should be done.

It will be used partitioning by quarters and every quarter will be partitioned by operations.

Range Partition by EVENT\_DT. The transactions are divided by quarters.

|  |  |  |
| --- | --- | --- |
| Quarter number | Start date | End date |
| 1 | 01-01 | 31-03 |
| 2 | 01-04 | 30-06 |
| 3 | 01-07 | 30-09 |
| 4 | 01-10 | 31-12 |

Hash partition of every quarter by channels. Number of sub partitions is 3, because transaction available in three operations.

The result table

|  |  |  |  |
| --- | --- | --- | --- |
| FCT\_TRANSACTIONS | Quarter 1 | Hash\_1 | facts of transactions |
| facts of transactions |
| facts of transactions |
| … |
| Hash\_2 | … |
| Hash\_3 | … |
| Quarter 2 | Hash\_1 | … |
| Hash\_2 | … |
| Hash\_3 | … |
| Quarter 3 | Hash\_1 | … |
| Hash\_2 | … |
| Hash\_3 | … |
| Quarter 4 | Hash\_1 | … |
| Hash\_2 | … |
| Hash\_3 | … |

* 1. **Strategy of Parallel execution**

Data Manipulation Language (DML) operations such as INSERT, UPDATE, and DELETE can be parallelized. Parallel execution can speed up large DML operations and is particularly advantageous in data warehousing environments where it's necessary to maintain large summary or historical tables.

The data warehouse contains several large tables that need periodic updating. To increase the rate of this process may be used parallel in DML operations. This method can help to save a lot of time to update big FACTs and DIMMENSIONs tables.

Parallel execution improves processing for queries requiring large table scans, joins, or partitioned index scans.

To store the data will often come queries for the implementation of which is time consuming. Parallel query should be used to perform SELECT operation to prevent a long waiting of a response from the database

When you create a table or an index using parallel DDL, each parallel slave process allocates space based on the table or index's storage clause.

Using parallel in DDL operations (create tables as select, create index, alter index rebuild) saving a lot of time during updating structure of our DWH and star levels.