GENERAL ASSUMPTIONS

Presented schema was designed to fulfill the required KPIs retrievals mentioned in the project statement. All the data which is not needed in the process has not been included. We assume that all the measure's values are calculated in the previous ETL proces.

TABLES

Table **PEOPLE** stores ids of all the maintenance people with their airports but also one value (ID = -1, Airport = "-") which is placed in LOGBOOK_DAILY_INFO table when we want to indicate that reporting person is a pilot. So when we want to obtain a sum of logbook reports for maintenance stuff only we need to add clause: WHERE REPORTER_ID >= 0 to the query. Thanks to this it's not needed to store information like REPORTER_TYPE.

Table **AIRCRAFTS_DAILY_INFO** represents all the things which happened for particular aircraft, in particular day (it contains all the measures which require AIRCRAFTS ans DATES dimensions).

Table **LOGBOOK_DAILY_INFO** represents daily count of logbook reports for particular aircraft, made by different people. Such solutions ensures that there will be zero or one row with REPORTER_ID = -1 indicating the count of pilots's reports in logbook and zero to multiple rows with REPORTER_ID >= 0 indicating count of reports made by different maintenance people. To use the table to obtain KPIs we need join with **AIRCRAFTS_DAILY_INFO** table - this is optimized by materialized view **MV_LOGBOOK_INFO**.

MATERIALIZED VIEWS

We don't consider number of data updates and number of KPIs retrievals but we assume that there will be a large number of data. Having this in mind we created materialized views to pre-compute some of the aggregations, arithmetical operations and pre-execute joins on tables.

We used ENABLE QUERY REWRITE to make Oracle's query optimizer able to use defined MVs for faster query responding.
We decided to use FAST REFRESH because we assume that we will have huge amounts of historical data and statistics calculated in past months which will not change really often (we will mostly have new data coming) so using complete refresh will be not optimal. Materialized view logs have been created to enable FAST REFRESH. We assume refreshing ON DEMAND because materialized can be updated daily or monthly depending on the requested data.

Materialized view **MV_LOGBOOK_INFO** has been created to optimize Report Rates (RR). We don't have specific materialized views for MRRs and PRRs because we think that the joins made on tables have the highest cost. To obtain the MRR or PRR only simple WHERE clause is needed which is not that expensive. This is better solution than maintaining separate materialized views for pilots and maintenance people.