# Question 3

About chosing the technology for supporting the system under the assignment, the main concern is to be aware there are two main different processes. The stations and its slots generate a continous stream of data which must be inmediately updated in the Databases, while on the other hand, anyone could need to demand a response for question 2 on any district, or any similar question.

The first process belongs to the OLTP class of database processes, and the second one is a typical analysis process from the OLAP side. "Question 1" is a simple query on the side of OLTP processes.

OLTP and OLAP processes feed on different time frames. Bike relocation process cares only for the current second or minute, while 'Question-2' process needs the hystorical register, past year, past 2 years, and so on. OLTP manages **high throughput** and **concurrency** of update/write and read queries scenarios while OLAP manages moving, filtering a analyzing big bulks of data under a regime of batch schedulings. 'Question-2' is an OLAP process and usual real-time update of slots in stations is a OLTP process.
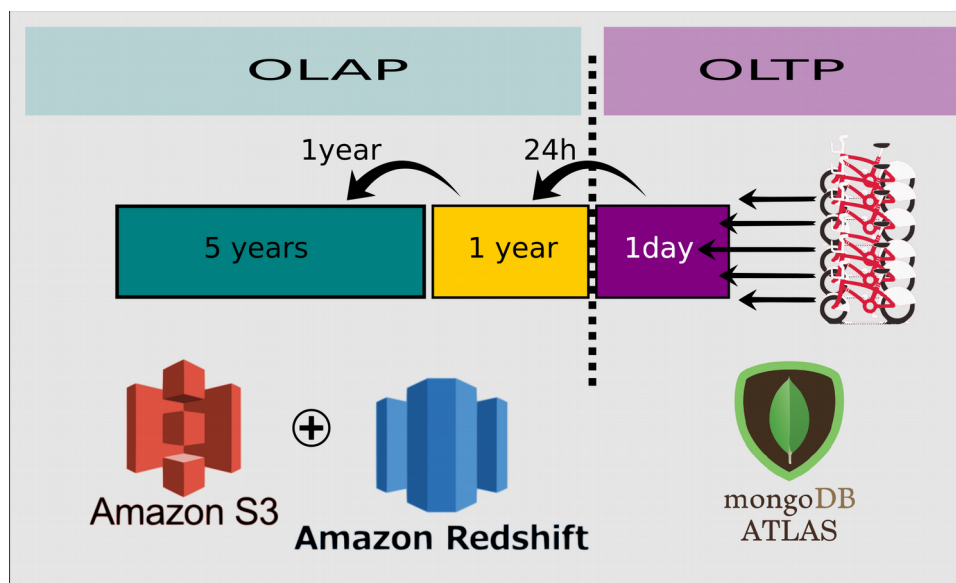
The proposal is to keep all data at cloud with **Amazon Web Services(AWS)**. It is an easy to work with platform for integration of big storage and data mining. AWS gives great flexibility for all type of options and tools are at your disposal.

For the OLAP part the adviced tool/platform is **REDSHIFT**.

For the OLTP part the advice would be to use RDS. But as RDS is based on RDBMS we would be forced to use SQL. And we would like to keep the same flavour of the Bicing API with JSON documents and all the MondoDB dependency from the scripts. The OLTP proposal changes in favor to **MongoDB Atlas**. MongoDB Atlas is the platform for MongoDB under AWS. As secondary virtue, Mongo eases data <u>sharding</u>, which is <u>crucial for scalability</u>, which is very much desired for many stations and many transactions writing and updating concurrently in the database.

Once a day, OLTP layer dumps its data at night to the OLAP layer servers. Just for adding one last turn of the screw, the OLAP layer could be managed as a hierarchical storage system. OLTP dumps to a first level which spans 1 year of history, which in turns feeds a deeper level which spans 5 year history of data.