

## Assumptions-

1. Cron job: Expect to receive the unprocessed files from the folder specified.
2. Input files from supplier:
  - Assuming that the supplier follows a contract in terms of what metadata they send. Need to have the specified keywords for product id and quantity to identify them in the metadata and accordingly parse the same, or we could have a set of keywords that the suppliers use and hence use this for our benefit.
  - Assuming that the metadata is always present in the first line
3. File size: The file which we are processing fits in memory. To tackle this we can have an implementation which processes file line by line or chunk of the file which fits in memory.
4. Data persistence: Assuming that we need to persist only 3 fields SupplierId, ProductId and Quantity. If this is not the case we need to go with a NoSQL database which will be used to store the unstructured data sent by the suppliers.
5. We can use respective ORMs to persist the SupplierInfo or write adaptors for easier mapping of objects to tables.

## Design-

1. Database: Defined an DBManager interface to keep database options open. Since, in our case we want to insert or update data and don't have to know what is the underlying database we are writing into. Let's say we go with SQL database, we just have to implement DBManager and so on. Thereby keeping the design accommodative of the changing databases.
2. Parser: With the Parser interface we can have multiple implementations of the parser based on the file types. Thereby the design is open to accommodate new additions.

## Schema-

Since the purpose of data is not known to us, we can broadly go with the following-

1. Analytics on the stored data or keeping all the information sent by the supplier, we can go with NoSQL database.
2. If we need transactional control over the data, we can go with SQL databases, where we can have the following schema-

if we need only 3 fields, we can have the suggested schema in the problem statement

Table: Product

\*ProductId, \*SupplierId, Quantity

primary\_key(ProductId, SupplierId)

**else**

Table: **Product**

\*ProductId(primary key), SupplierId(foreign key), Quantity, more product info fields

Table: **Supplier**

\*SupplierId(primary key), more supplier info fields