Batch Processing

1. Bulk Processing
2. Banking: Bulk money transfer at night.
3. E-Commerce/Retail: Uploading items in bulk at night.

Spring Batch

1. It is a lightweight batch framework based on spring.
2. It is used for high volume batch processing.
3. Various I/O options like JSON, XML, SQL databases.
4. Easy to manage with start/stop options.
5. Retry and skip mechanism.

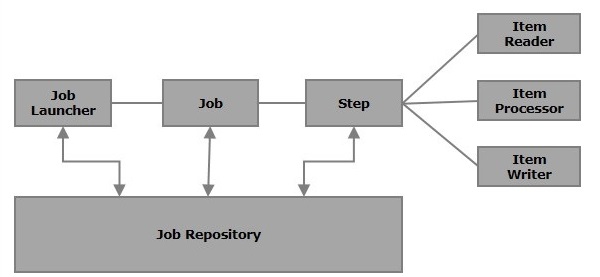
What is job and step?

* Job can be online transaction and step is like multiple steps in job like deduct money from the sender then add money to receiver.

Step types in Spring Batch

1. **Tasklet Step**: Online bulk transaction at night.
2. **Chunk-Oriented Step**: Processing chunk of CSV at a time. Like 5 rows at a time. The flow here is Item Reader -> Item Processor -> Item Writter

**Spring Batch Architecture**



We must add Spring Batch dependency in pom.xml

We must add @EnableBatchProcessing in main class

We must have at least one sql database connection otherwise spring batch application will not run.

We can use command line arguments or parameters to uniquely run one job

One can use RunIdIncrementer class but for that for the first time we have to pass some unique job parameters

ItemProcessor is not a mandatory step.

We can have chunk-oriented step and tasklet step in the same job

By default, spring batch will run all the jobs on the boot-up which have @Bean annotation. For local running it is helpful but for not for production, so we must schedule our spring batch job and we must run job manually on demand or in a scheduled way.

To prevent auto start we can add this config to property file **spring.batch.job.enabled: false**

We can either start via rest call or via spring schedular

When we manually stop any job then that job will be stop but the chunk whatever it was working that will be completed then gracefully it will be stopped.

Item readers can be anything:

1. Flat file (CSV)
2. JSON
3. XML
4. Database
5. Rest Api

For each of these source spring batch provides different methods

Flat file item reader:

1. Source location of CSV
2. Line Mapper:
3. Line Tokenizer (,)
4. Bean Mapper (converting from one line to object)

Flat file item reader CSV

If we change the order of the name here: **delimitedLineTokenizer.setNames("id", "first\_name", "last\_name", "email");**

then the value of the fields of the POJO class will also be changed.

If we remove any name from here, we will get parse exception.

We can change the delimiter from comma to pipe separator.

We can also provide the csv path from job arguments; we can hold that from **@Value(“#{jobParameters[‘inputFile’]}”)**

If we mark it **flatFileItemReader()** as a **@Bean** then we must add **@StepScope**

Also we can use the **classPathResource** if the file is inside of **“src/main/resources”**