

CLAIMS

What is claimed is

1. Method for processing data for database modification, comprising:
 - 5 - receiving a set of data,
 - performing a processing chain comprising a plurality of consecutive jobs to transform the set of data into transformed data,
 - modifying a production database with respect to the transformed data,further comprising the steps of:
 - 10 - setting a target processing time for the performance of the consecutive jobs,
 before a launch of a first job, applying an original configuration as current configuration defining a parallelization level for each of the consecutive jobs,
 upon an actual remaining processing time being out of a range of
15 acceptable remaining processing times, applying an adapted configuration as new current configuration defining an adapted parallelization level for each of the jobs remaining in the processing chain, said adapted configuration differing from the current configuration.
2. The method of the preceding claim comprising:
 - 20 before a launch of all further jobs after the first job, upon an actual remaining processing time being out of a range of acceptable remaining processing times, applying an adapted configuration as new current configuration defining an adapted parallelization level for each of the jobs remaining in the processing chain, said adapted configuration differing from the current configuration.
- 25 3. The method according to any of the preceding claims wherein the range of acceptable remaining processing times is defined as a range of times under a determined proportion of the target processing time.
4. The method according to any of the preceding claims wherein the original configuration and the adapted configuration are determined on the basis of information
30 comprising historical data and constraint data.
5. The method of the preceding claim wherein the historical data comprise information on previous executions of processing chains for sets of data of a same type as the set of data.

6. The method of the preceding claim comprising the step of creating a configuration table comprising, for several ranges of data volume, a configuration definition section including at least one configuration definition specifying a number of parallel instances to be applied for each job.

5 7. The method of the preceding claim comprising, upon the actual remaining processing time being under the range of acceptable remaining processing times, choosing the adapted processing configuration among at least one configuration of the configuration table for which the number of parallel instances to be applied to the remaining jobs is higher than the number of instances to be applied to the remaining
10 jobs of the current configuration.

8. The method of claim 6 comprising, upon the actual remaining processing time being over the range of acceptable remaining processing times, choosing the adapted processing configuration among at least one configuration of the configuration table for which the number of parallel instances to be applied to the remaining jobs is
15 lower than the number of instances to be applied to the remaining jobs of the current configuration.

9. The method according to any of the preceding claims comprising:

- receiving at least another set of data,
- detecting at least one dependent job of a processing chain of the
20 another set of data which depends from at least one given job of the processing chain of the set of data,
- scheduling a launch of the dependent job after the completion of the one given job.

10. The method of any of the preceding claims comprising using a set of data
25 including data entities each describing one component of a fare definition.

11. The method of the preceding claim wherein the component is selected among fare and fare rules and routings.

12. The method according to any of the two preceding claims wherein the processing chain comprises:

- 30 - A first job of conversion of the set of data into a set of data converted in a format of an internal structure,
- A second job of construction of at least one preliminary table based on the converted set of data,
- A third job of update of a reference database based on the at least one
35 preliminary table,

- A fourth job of load of the update of the reference database into a production database.

13. The method of the preceding claim wherein the first job is using as input a file containing the set of data.

5 14. The method of any of the two preceding claims wherein the fourth job is launched for at least one data entity for which the third job is completed even if the third job is not completed for all the data entities.

10 15. The method of method of any of the three preceding claims comprising receiving a set of data for each of components of fare definitions and performing parallel processing chains for said components.

16. The method of any of the preceding claims performed as a result of execution of computer software by at the least one data processor, the computer software being stored in a non-transitory computer-readable memory medium.

15 17. The method of any of the preceding claims comprising sending the set of data from a provider system.

18. The method of any of the preceding claims comprising accessing data of the production database from and end user device via a search engine.

19. A system for processing data for database modification, comprising means configured to execute the method of any of the preceding claims.

20 20. System of the preceding claim comprising a resource allocator comprising means for:

25 setting a target processing time for the performance of the consecutive jobs,
 before a launch of a first job, applying an original configuration as current configuration defining a parallelization level for each of the consecutive jobs,
 upon an actual remaining processing time being out of a range of acceptable remaining processing times, applying an adapted configuration as new current configuration defining an adapted parallelization level for each of the jobs remaining in the processing chain, said adapted configuration differing from the current configuration.

30 21. System of the preceding claim comprising a job scheduler having means for triggering the launch of the jobs.

35 22. System of the preceding claim wherein the job scheduler comprises means for, upon receipt of at least another set of data, detecting at least one dependent job of a processing chain of the another set of data which depends from at least one given job of the processing chain of the set of data, and means for scheduling a launch of the dependent job after the completion of the one given job.

23. System of the preceding claim, comprising computer software stored in a non-transitory computer-readable memory medium that is executed by at least one data processor that comprises part of the system.

5 24. Computer program product stored in a non-transitory computer-readable memory medium and comprising instructions adapted to perform the method of any of claims 1 to 18.

CLAIMS

1. A method for processing data for database modification, the method comprising:

5 receiving a first set of data;

performing a processing chain comprising a plurality of consecutive jobs to transform the first set of data into transformed data;

modifying a database with respect to the transformed data;

setting a target processing time for performance of the consecutive jobs;

10 before the consecutive jobs are launched, applying a first configuration defining a first number of parallel instances for each of the consecutive jobs; and

if an actual remaining processing time for the consecutive jobs uncompleted in the processing chain is outside of a range of acceptable remaining processing times relative to the target processing time, applying a second configuration defining a

15 second number of parallel instances for each of the consecutive jobs uncompleted in the processing chain and including the first number of parallel instances for each of the consecutive jobs completed in the processing chain, the second number of parallel instances in the second configuration for at least one of the consecutive jobs uncompleted in the processing chain differing from the first number of parallel
20 instances in the first configuration,

wherein the first configuration and the second configuration are determined based upon information comprising historical data and constraint data, and the historical data comprises information on previous executions of processing chains for sets of data of a same type as the first set of data.

25

2. The method of claim 1 wherein the range of acceptable remaining processing times is defined as a range of times under a determined proportion of the target processing time.

30

3. The method of claim 1 comprising:

creating a configuration table comprising, for several ranges of data volume, a configuration definition section including a plurality of configurations each specifying
5 a number of parallel instances to be applied for each job,
wherein the first configuration is among the plurality of configurations.

4. The method of claim 3 further comprising: storing the second configuration in the configuration table.

10

5. The method of claim 1 wherein the second configuration is applied if the actual remaining processing time is under the range of acceptable remaining processing times, and the second number of parallel instances to be applied to the remaining jobs is higher than the first number of parallel instances to be applied to
15 the remaining jobs.

6. The method of claim 1 wherein the second configuration is applied if the actual remaining processing time is over the range of acceptable remaining processing times, and the second number of parallel instances to be applied to the remaining jobs is lower than the first number of parallel instances to be applied to the
20 remaining jobs.

7. The method of claim 1, wherein the first set of data comprises data entities, each data entity describing a component of a fare definition of a travel product.

25

8. The method of claim 7 wherein the component is a fare, a fare rules, or a routing.

9. The method of claim 7 wherein the processing chain comprises:

30 a first job of conversion of the first set of data into a first set of data converted in a format of an internal structure,

a second job of construction of at least one preliminary table based on the converted first set of data,

a third job of update of a reference database based on the at least one preliminary table, and

5 a fourth job of load of the update of the reference database into a production database.

10. The method of claim 9 wherein the first job is using as input a file containing the first set of data.

10

11. The method of claim 9 wherein the fourth job is launched for at least one data entity for which the third job is completed even if the third job is not completed for all the data entities.

15 12. The method of claim 9 comprising: receiving the first set of data for each of the data entities of components of fare definitions; and performing parallel processing chains for said components the data entities.

20 13. The method of claim 1 comprising: sending the first set of data from a provider system.

14. The method of claim 1 comprising: accessing data of the production database from an end user device via a search engine.

25 15. The method of claim 1 further comprising: upon receipt of a second set of data, detecting a dependent job of a processing chain of the second set of data that requires input data relying on results generated by a given job of the processing chain of the first set of data; and scheduling a launch of the dependent job after the completion of the given job.

30

16. The method of claim 1 wherein applying the first configuration defining the first number of parallel instances for each of the consecutive jobs comprises: selecting

the first configuration based upon the data in the set numbering between a minimum and a maximum.

17. The method of claim 16 wherein the second number of parallel instances for
5 each of the consecutive jobs uncompleted in the processing chain is obtained from a third configuration that is applied if the data in the first set number is greater than the maximum.

18. The method of claim 16 wherein the second number of parallel instances for
10 each of the consecutive jobs uncompleted in the processing chain is obtained from a third configuration that is applied if the data in the first set number is less than the minimum.

19. A system comprising:
15 at least one data processor; and program code configured upon execution by the at least one processor to process data for database modification by:
receiving a first set of data;
performing a processing chain comprising a plurality of consecutive jobs to transform the first set of data into transformed data;
20 modifying a database with respect to the transformed data;
setting a target processing time for performance of the consecutive jobs;
before the consecutive jobs are launched, applying a first configuration defining a first number of parallel instances for each of the consecutive jobs; and
if an actual remaining processing time for the consecutive jobs uncompleted in the
25 processing chain is outside of a range of acceptable remaining processing times relative to the target processing time, applying a second configuration defining a second number of parallel instances for each of the consecutive jobs uncompleted in the processing chain and including the first number of parallel instances for each of the consecutive jobs completed in the processing chain, the second number of
30 parallel instances in the second configuration for at least one of the consecutive jobs uncompleted in the processing chain differing from the first number of parallel instances in the first configuration,

wherein the first configuration and the second configuration are determined based upon information comprising historical data and constraint data, and the historical data comprises information on previous executions of processing chains for sets of data of a same type as the first set of data.

5

20. The system of claim 19, wherein at least one of the first and second sets of data comprise data entities, each data entity describing a component of a fare definition of a travel product.

10 21. The system of claim 19 wherein the program code configured upon execution by the at least one processor to process data for database modification by:

upon receipt of a second set of data, detecting a dependent job of a processing chain of the second set of data that requires input data relying on results generated by a given job of the processing chain of the first set of data; and

15 scheduling a launch of the dependent job after the completion of the given job.

22. The system of claim 19 wherein the program code configured upon execution by the at least one processor to process data for database modification by applying the first configuration defining the first number of parallel instances for each of the
20 consecutive jobs comprises:

program code configured upon execution by the at least one processor to process data for database modification by selecting the first configuration based upon the data in the set numbering between a minimum and a maximum.

25 23. The system of claim 22 wherein the second number of parallel instances for each of the consecutive jobs uncompleted in the processing chain is obtained from a third configuration that is applied if the data in the first set number is greater than the maximum.

30 24. The system of claim 22 wherein the second number of parallel instances for each of the consecutive jobs uncompleted in the processing chain is obtained from a third configuration that is applied if the data in the first set number is less than the

minimum.

25. A computer program product comprising:

a non-transitory computer-readable storage medium; and a computer program
5 stored on the storage medium, the computer program comprising instructions that, when executed on a computer, cause the computer to process data for database modification by:

receiving a first set of data;

performing a processing chain comprising a plurality of consecutive jobs to
10 transform the first set of data into transformed data;

modifying a database with respect to the transformed data;

setting a target processing time for performance of the consecutive jobs;
before the consecutive jobs are launched, applying a first configuration defining a
first number of parallel instances for each of the consecutive jobs; and

15 if an actual remaining processing time for the consecutive jobs uncompleted in the processing chain is outside of a range of acceptable remaining processing times relative to the target processing time, applying a second configuration defining a second number of parallel instances for each of the consecutive jobs uncompleted in the processing chain and including the first number of parallel instances for each of
20 the consecutive jobs completed in the processing chain, the second number of parallel instances in the second configuration for at least one of the consecutive jobs uncompleted in the processing chain differing from the first number of parallel instances in the first configuration,

wherein the first configuration and the second configuration are determined based
25 upon information comprising historical data and constraint data, and the historical data comprises information on previous executions of processing chains for sets of data of a same type as the first set of data.

26. The computer program product of claim 25, wherein at least one of the first
30 and second sets of data comprise data entities, each data entity describing a component of a fare definition of a travel product.

27. The computer program product of claim 25 wherein the instructions that, when executed on the computer, further cause the computer to process data for database modification by:

- 5 upon receipt of a second set of data, detecting a dependent job of a processing chain of the second set of data that requires input data relying on results generated by a given job of the processing chain of the first set of data; and
 scheduling a launch of the dependent job after the completion of the given job.

10 28. The computer program product of claim 25 wherein the instructions that, when executed on the computer, cause the computer to process data for database modification by applying the first configuration defining the first number of parallel instances for each of the consecutive jobs comprises:

- instructions configured upon execution by the at least one processor to
15 process data for database modification by selecting the first configuration based upon the data in the set numbering between a minimum and a maximum.

29. The computer program product of claim 28 wherein the second number of parallel instances for each of the consecutive jobs uncompleted in the processing
20 chain is obtained from a third configuration that is applied if the data in the first set number is greater than the maximum.

30. The computer program product of claim 28 wherein the second number of parallel instances for each of the consecutive jobs uncompleted in the processing
25 chain is obtained from a third configuration that is applied if the data in the first set number is less than the minimum.