

METHOD AND APPARATUS OF DATA PROCRESSING FOR ONLINE ANALYTICAL PROCESSING

This application claims priority to Chinese Application No. 201810160009.7, filed
5 on February 26, 2018, which is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

Embodiments of the present invention relate to the field of data processing technologies, and in particular to a method and an apparatus of data processing for online analytical processing.

10 BACKGROUND

Nowadays the internet has massive data, and a rapid upgrade of wireless networks, a wide use of terminal equipments and an online development of insurance businesses make it convenient for users to perform a check, a purchase, an edit and other operations of the insurance. In order to perform a data analysis and decision for the data, the user and
15 the insurance, these data are stored electronically, which makes a content and a quantity of an insurance data become more and more.

An online insurance business data is mainly stored in a business database, and an online database mainly uses an OLTP (on-line transaction processing) mode to adapt to rapid operations such as creating, modifying, editing and deleting, namely, to support a
20 rapid response to transactions.

In order to maximize the value of the data , data analysts, business staffers and decision makers analyze these massive insurance data and find out the characteristics of users, customer groups and insurance businesses to provide a data basis for a data operation and a decision , and this data analysis mode realizes a rapid and
25 multidimensional analytical search of the data in an OLAP (online analytical processing system) way.

Because of differences of use between online transaction data and offline analysis data, the data needs to be transferred from OLTP to OLAP mode. An existing data processing way is concatenating the data with different types according to information
30 such as a policy number ID, a user and the like, and then performing a multidimensional

analysis of the data through the OLAP.

A traditional internet data concatenation solution is to perform a batch process to the data according to a day, a week or a month, and then perform a multidimensional analysis of data. But this solution is not suitable for the insurance data with massive data, as compared with a traditional data, an insurance finance data will be modified within a certain time, and this modification of the data also needs to be reflected in a data warehouse of the OLAP, that is, a concatenation process or a union process to the modification of the data and the data in the data warehouse of the OLAP before the modification needs to be performed by backtracking. More unnecessary calculation resources will be consumed through the batch process when the quantity of the modification of the data is large.

SUMMARY

In view of this, embodiments of the present invention provide a method and an apparatus of data processing for online analytical processing, which can realize a rapid concatenation of data, and reduce unnecessary data concatenating calculation amounts when business data with massive data changes frequently. These technical solutions are as follows.

In a first aspect, a method of data processing for online analytical processing is provided, and the method includes:

performing a partition division on an incremental data of an online analytical processing system according to a selected partition field, and obtaining a plurality of incremental data partitions;

performing a partition division on a stock data of the online analytical processing system according to the selected partition field, and obtaining a plurality of stock data partitions; and

integrating data in the incremental data partitions and data in the stock data partitions which have a same partition identity.

Combining with the first aspect, in a first possible embodiment, the integrating data in the incremental data partitions and data in the stock data partitions which have a same partition identity includes:

performing a concatenation operation and/or a union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions.

Combining with the first possible embodiment of the first aspect, in a second possible embodiment, the method further includes:

- 5 constructing a data index of the plurality of the stock data partitions; and performing the concatenation operation and/or the union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions.

- 10 Combining with any one possible implementation of the first aspect to the second possible implementation of the first aspect, in a third possible embodiment, the partition field includes anyone or any combinations of a time field, a business field and a custom field.

- 15 Combining with the second possible embodiment of the first aspect, in a fourth possible embodiment, the constructing a data index of the plurality of the stock data partitions includes:

- constructing the data index in a manner of a binary tree or a hash.

- 20 Combining with the second possible embodiment of the first aspect, in a fifth possible embodiment, the performing the concatenation operation and/or the union operation on the data in each of incremental data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions includes:

- searching for the data in each of the incremental data partitions according to the data index, in the corresponding stock data partitions; and

- 25 performing the concatenation operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is successful.

- 30 Combining with the second possible embodiment of the first aspect or the fifth possible embodiment of the first aspect, in a sixth possible embodiment, the performing the concatenation operation and/or the union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions includes:

- searching for the data in each of the incremental data partitions according to the data index, in the corresponding stock data partitions; and

performing the union operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is unsuccessful.

Combining with any one possible implementation of the first aspect to the sixth possible implementation of the first aspect, in a seventh possible embodiment, after the
5 integrating data in the incremental data partitions and data in the stock data partitions which have a same partition identity, the method further includes:

exporting an integrated data to a specific location through transforming and loading in order to perform a data analysis of the online analytical processing system; and/or

outputting prompt information, which is used for indicating that the data is updated,
10 to a downstream data processing unit.

In a second aspect, an apparatus of data processing for online analytical processing is provided, and the apparatus includes:

a partition divisioning module, configured to perform a partition division on an incremental data of an online analytical processing system according to a selected
15 partition field, and obtain a plurality of incremental data partitions;

the partition divisioning module, further configured to perform a partition division on a stock data of the online analytical processing system according to the selected partition field, and obtain a plurality of stock data partitions; and

a data processing module, configured to integrate data in the incremental data
20 partitions and data in the stock data partitions which have a same partition identity.

Combining with the second aspect, in a first possible embodiment, the data processing module is further configured to perform a concatenation operation and/or a union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions.

25 Combining with the first possible embodiment of the second aspect, in a second possible embodiment, the apparatus further includes:

a data index constructing module, configured to construct a data index of the plurality of the stock data partitions; and the data processing module performs the concatenation operation and/or the union operation on the data in each of the incremental
30 data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions.

Combining with any one possible implementation of the second aspect to the second possible implementation of the second aspect, in a third possible embodiment, the

partition field includes anyone or any combinations of a time field, a business field and a custom field.

Combining with the second possible embodiment of the second aspect, in a fourth possible embodiment, the data index constructing module is configured specifically to
5 construct the data index in a manner of a binary tree or a hash.

Combining with the second possible embodiment of the second aspect, in a fifth possible embodiment, the data processing module includes:

 a search submodule, configured to search for the data in each of the incremental data partitions according to the data index, in the corresponding stock data partitions; and
10 a concatenation operation submodule, configured to perform the concatenation operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is successful.

Combining with the fifth possible embodiment of the second aspect, in a sixth possible embodiment, the data processing module further comprises:

15 a union operation submodule, configured to perform the union operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is unsuccessful.

Combining with any one possible implementation of the second aspect to the sixth possible implementation of the second aspect, in a seventh possible embodiment, the
20 apparatus further includes:

 a data exporting module, configured to export an integrated data to a specific location through transforming and loading in order to perform a data analysis of the online analytical processing system.

Combining with any one possible implementation of the second aspect to the seventh possible implementation of the second aspect, in an eighth possible embodiment, the
25 apparatus further includes:

 an information outputting module, configured to output prompt information, which is used for indicating that the data is updated, to a downstream data processing unit after integrating.

30 The embodiments of the present invention provide a method and an apparatus of data processing for online analytical processing, by performing respectively a partition division on an incremental data and a stock data of an online analytical processing system and integrating data in the incremental data partitions and data in the stock data partitions

which have a same partition identity, so that a data integration process can be performed rapidly, and unnecessary operations to all data is also avoided and the calculation amounts of the data are reduced greatly. In addition, the data is integrated rapidly, which is helpful for business staffers to perform data and business explorations based on an insurance data
5 with a complete dimensionality.

BRIEF DESCRIPTION OF DRAWINGS

In order to illustrate technical solutions in the embodiments of the present invention more clearly, the drawings used in the description of the embodiments will be briefly introduced below. Obviously, the drawings described below are only some embodiments
10 of the present invention. For those skilled in the art, other drawings may be obtained according to these drawings without any creative work.

FIG. 1 is a flowchart of a method of data processing for online analytical processing according to an embodiment of the present invention;

FIG. 2 is a flowchart of a method of data processing for online analytical processing
15 according to an embodiment of the present invention;

FIG. 3 is a result of a data partition and a result of a data index construction to the partition according to an embodiment of the present invention; and

FIG. 4 is a block diagram of an apparatus of data processing for online analytical processing according to an embodiment of the present invention.

DETAILED DESCRIPTION

20

In order to more clearly illustrate the purpose, technical solution and advantages of the present invention hereinafter specific embodiments of the present invention will be clearly completely described with reference to the appended drawings. Apparently, the embodiments described are only a part but not all of the embodiments of the present
25 invention. For those skilled in the art, all other embodiments obtained based on the embodiments of the present invention without creative work should be embraced within the protection scope of the present invention.

An embodiment of the present invention provides a method of data processing for online analytical processing, and the method may be applied to a scenario where an
30 insurance data with a large amount of data changes frequently under the , realize a rapid

concatenation of the data, and reduce the unnecessary data concatenating calculation amounts. In addition, the method provided by the embodiment of the present invention also may be applied to other business data with a large amount of data, and the present invention does not limit specific application scenarios.

5 FIG. 1 is a flowchart of a method of data processing for online analytical processing according to an embodiment of the present invention, and referring to FIG. 1, the method includes the following steps.

101: a partition division is performed on an incremental data of an online analytical processing system according to a selected partition field, and a plurality of incremental
10 data partitions are obtained.

102: a partition division is performed on a stock data of the online analytical processing system according to the selected partition field, and a plurality of stock data partitions are obtained.

103: data in the incremental data partitions and data in the stock data partitions
15 which have a same partition identity are integrated.

FIG. 2 is a flowchart of a method of data processing for online analytical processing according to an embodiment of the present invention, and referring to FIG. 2, the method includes the following steps.

201: a partition division is performed on an incremental data of an online analytical
20 processing system according to a selected partition field, and a plurality of incremental data partitions are obtained.

Wherein, since a same data type, such as the transaction ID, the unique ID and so on, is usually used for concatenating the incremental data with a stock data. And changes of the stock data is less and less over time, thus a field that does not change after creating
25 data may be selected as a partition field to perform the partition division on the incremental data and the stock data, wherein the selected partition field may include anyone or any combinations of a time field, a business field and a custom field. The time field may be a creation time of a transaction data, and the business field may be a field such as a channel or a region and the like.

30 Specifically, step 101 may include followings.

An ETL (Extract-Transform-Load) process, including data extracting and data transforming, is performed to the incremental data that is newly added in the online analytical processing system, to transform the incremental data into a storage mode

required for an online analytical processing; and

the field that does not change after creating data is selected as the partition field, then the partition division on the incremental data that is performed with the ETL process is performed according to the partition field, and the plurality of the incremental data partitions are obtained.

In addition, when the partition divisions to the incremental data and the stock data are performed, more options of partition modes can be supported through abstracting a data partition. For example, when the time field is used as the partition field, a time partition may be abstracted and be adjusted to a more detailed partition organization according to a periodic processing mode of a data task, namely: if a processing period of data is a year, a month, a week, a day, and so a corresponding partition period of the data is also the year, the month, the week, the day.

In addition, when the partition divisions to the incremental data and the stock data are performed, a more fine-grained partition mode may be also added, and detailed grained partition fields are integrated according to actual needs of a business. For example, if the transaction data respectively corresponding to a channel and a region of the actual business do not overlap each other, then the time field + the business field [channel, region] may be used as the partition field to perform the partition divisions to the incremental data and the stock data respectively.

202: a partition division is performed on a stock data of the online analytical processing system according to the selected partition field and a plurality of stock data partitions are obtained, and a data index of the plurality of the stock data partitions is constructed.

Wherein the stock data may be stored on a big data platform in order to expand horizontally rapidly, for example, a big data may be stored through Hadoop.

Wherein the partition field that is the same as the partition field of the partition division of the incremental data, is used for performing the partition division on the stock data, and the plurality of the stock data partitions are obtained.

Wherein the process of constructing a data index of the plurality of the stock data partitions may include:

constructing a data index table for each of the stock data partitions in a manner of a binary tree, wherein the data index table has a data structure of the binary tree, and includes a mapping relationship between a data identity and the stock data.

The binary tree is a special tree, and in which each node has a maximum number of two child nodes that generally called a left child node and a right child node, and child trees of the binary tree include right child trees and left child trees, and the order of the child trees can not be arbitrarily reversed. The binary tree is defined recursively.

5 Besides, the data index of the plurality of the stock data partitions may be constructed in the manner of a hash.

The embodiment of the present invention does not limit the specific data index constructing process.

10 In the embodiment of the present invention, an indexed mode of Hbase may be used for performing position and update operations of data, and Mysql databases may be also used for performing a constructing of the index and an update of of the data, and a small local file database may be also used for performing the constructing of the index and the update of the data, such as levelDB.

15 203: corresponding relationships between the plurality of the incremental data partitions and the plurality of the stock data partitions are established according to whether they have same partition identities.

Wherein the plurality of the incremental data partitions and the plurality of the stock data partitions have their own partition identities respectively;

20 specifically, the embodiment of the present invention does not limit specific establishing processes.

FIG. 3 shows a result of a data partition and a result of a data index construction to the data partition in an embodiment of the present invention. In FIG. 3, a constructing time is used as a partition field, according to which partition divisions to an incremental data and a stock data of an online analytical processing system are performed respectively, and a plurality of incremental data partitions and a plurality of stock data partitions are obtained, where the plurality of the incremental data partitions and the plurality of the stock data partitions all include {day}, {day-1}, ... {day-n}, {day-n-1} A corresponding relationship between the incremental data partitions and the stock data partitions which have a same partition identify may be established; in addition, as can be seen in FIG. 3, the data index in each partition of the stock data corresponds to each partition one by one.

In the embodiment of the present invention, by establishing the corresponding relationships between the plurality of the incremental data partitions and the plurality of

the stock data partitions according to whether they have same partition identities, it is convenient for a concatenation or a union process to data in the same partition subsequently, so that an unnecessary data concatenating try is reduced, and an unnecessary data concatenating calculation amount is avoided.

5 204: a concatenation operation and/or a union operation on data in each of the incremental data partitions and data in corresponding stock data partitions is performed based on the data index of the plurality of the stock data partitions.

Specially, the process may include:

Step a: a search in the stock data partitions corresponding to each of the incremental
10 data partitions, is performed for the data in each of the incremental data partitions according to the data index.

Specially, the process may include:

a data identity of the data in each of the incremental data partitions is acquired, and the search to the data identity in a data index table is performed. If the data identity is
15 the data index table, the search to the stock data corresponding to the data identity in the data index table is successful.

Step b: if the search result is successful, the concatenation operation on the incremental data and associated stock data is performed, and if the search result is unsuccessful, the union operation on the incremental data is performed.

20 Exemplary, a transaction data is generated by user1 on October 10, 2017, and corresponding transaction data ID is ID_user1 and a constructing time is October 10, 2017. The transaction data is stored in a data warehouse of an online analytical processing system as a stock data, and a partition to the stock data is performed through a time field, so the stock data is divided into a partition of October 10, 2017. If the transaction data of
25 ID_user1 is modified by user1 on December 15, 2017 because of a business reason, and a modified content is any other data except for the transaction data ID, such as modified contact information, added identity information and the like, the transaction data at this time is transferred to the online analytical processing system as an incremental data, the partition to the incremental data is performed through the time field, and so the
30 incremental data is divided into the partition of October 10, 2017. Obviously, the transaction data of ID_user1 can be found in the partition of October 10, 2017, and then a data concatenation process is performed, namely, the data concatenation operation on data in the partition of October 10, 2017 of a modified data and data in the partition of October

10, 2017 of a historical stock data is performed. The specific concatenation is processed according to the transaction data ID which is ID_user1, and the modified contact information and the added identity information are stored in a historical data by the data concatenation process.

5 Optionally, after step 204, the method provided by the embodiment of the present invention further includes:

 exporting data after performing a concatenation operation and/or a union operation on a specific location through transforming and loading, in order to perform a data analysis of the online analytical processing system; and/or

10 outputting prompt information, which is used for indicating that the data is updated, to a downstream data processing unit.

 The method of data processing for online analytical processing provided by the embodiment of the present invention is used for, in the actual application when a data processing is performed for OLAP, performing a partition organization to the data, which
15 may be in a common partition organization way according to a constructing time, and when HBase is used as a storage scheme for the stock data, performing a sort and a search on an index, which may be in a rowkey way, and thus performing a rapid location to the data. In addition, because Hive has characteristics of suiting SQL and supporting a plurality of search engines, Hive may be used as a basic storage of all historical data, and
20 thus based on Hive, it is made easier to perform a multidimensional OLAP analysis. In addition, for an update of each batch of data, partitions are processed one by one, and when each partition is processed, the data that needs to be updated is located and read by rowkey, and the data is updated and written back to HBase at the same time. Finally, the partition data changed is exported to Hive directly , and since an export is not related to
25 update and union operations of the data, the export is performed relatively rapidly.

 The method of data processing for online analytical processing provided by the embodiment of the present invention is by performing partition divisions to an incremental data and a stock data of an online analytical processing system respectively, and constructing a data index to a partition, and performing a concatenation or a union
30 process in the same partition. Since data is searched and corresponding data is updated through an index, only data that needs to be updated is searched and updated, so that the processes of the concatenation or the union of the data may be performed rapidly, and an unnecessary operation in all data is also avoided, which simplifies the calculation amount

of the data greatly. In addition, by achieving to concatenate the data rapidly, it is helpful for business staffers to perform explorations of the data and the business based on an insurance data in a complete dimensionality. The method provided by the embodiment of the present invention has a great advantage in an insurance finance data scenario, especially when there is much data to be constructed and changed frequently in an insurance scenario, the advantages of the rapid concatenation and union operations of the data are more obvious.

FIG. 4 is a block diagram of an apparatus of data processing for online analytical processing according to an embodiment of the present invention, and referring to FIG. 4, the apparatus 3 includes:

a partition divisioning module 31, configured to perform a partition division on an incremental data of an online analytical processing system according to a selected partition field, and obtain a plurality of incremental data partitions; and further configured to perform a partition division on a stock data of the online analytical processing system according to the selected partition field, and obtain a plurality of stock data partitions; and

a data index constructing module 32, configured to construct a data index of the plurality of the stock data partitions;

a corresponding relationship establishing module 33, configured to establish a corresponding relationship between the plurality of the incremental data partitions and the plurality of the stock data partitions according to whether they have same partition identities;

a data processing module 34, configured to perform a concatenation operation and/or a union operation on data in each of the incremental data partitions and data in the corresponding stock data partitions.

Optionally, the partition field includes anyone or any combinations of a time field, a business field and a custom field.

Optionally, the data index constructing module 32 is specifically configured to: construct the data index in a manner of a binary tree or a hash.

Optionally, the data processing module 34 includes:

a search submodule 341, configured to search for the data in each of the incremental data partitions according to the data index, in the corresponding stock data partitions;

a concatenation operation submodule 342, configured to perform the concatenation operation on the incremental data and associated stock data when the search performed by

the search submodule 341 is successful; and

a union operation submodule 343, configured to perform the union operation on the incremental data when the search performed by the search submodule 341 is unsuccessful.

5 Optionally, the apparatus 3 further includes:

a data exporting module 35, configured to export data after performing the concatenation operation and/or the union operation on a specific location through transforming and loading in order to perform a data analysis of the online analytical processing system; and

10 an information outputting module 36, configured to output prompt information, which is used for indicating that the data is updated, to a downstream data processing unit.

The embodiment of the present invention provides an apparatus of data processing for online analytical processing. The apparatus performs a partition division on an incremental data and a stock data of an online analytical processing system respectively,
15 constructs a data index to a partition, and performs a concatenation or a union process in the same partition. Since the data is searched and corresponded data is updated through an index, only data that needs to be updated is searched and updated, so that the process the concatenation or the union of the data may be performed rapidly, and an unnecessary
20 operation in all data is also avoided, which simplifies the calculation amount of the data greatly. In addition, by achieving to concatenate the data rapidly, it is helpful for business staffers to perform explorations of the data and the business based on an insurance data in a complete dimensionality. The apparatus provided by the embodiment of the present invention has a great advantage in an insurance finance data scenario, especially when
25 there is much data to be constructed and changed frequently in an insurance scenario, the advantages of the rapid concatenation and union operations of the data are more obvious.

All of the above optional technical solutions may be combined in any way to form optional embodiments of the present invention, and details are not repeated herein.

It should be noted that when the apparatus of data processing for online analytical
30 processing provided by the above mentioned embodiments performs the method of data processing for online analytical processing, only the division of each functional module above is illustrated. In practical application, the above functions may be distributed by different functional modules to complete according to requirements. The internal structure

of the apparatus is divided into different functional modules to perform all or part of functions described above. In addition, the apparatus of data processing for online analytical processing and the method of data processing for online analytical processing provided by the above embodiments are in the same concept, and the specific
5 implementation process is described in details in the embodiments of method, and details are not described herein.

Those skilled in the art can understand that all or part of steps of implementing the above embodiments may be completed by hardware or the relevant hardware instructed by the program. The program may be stored in a computer readable storage medium, and
10 the storage medium mentioned above may be a read only memory, a magnetic disk, or an optical disc or the like.

The above embodiments are only the preferable embodiments of the present invention and are not intended to limit the present invention. Any modifications, equivalents substitutions, improvements, etc. made within the spirit and principle of the
15 present invention should be embraced within the protection scope of the present invention.

What is claimed is:

1. A method of data processing for online analytical processing, comprising:
performing a partition division on an incremental data of an online analytical processing system according to a selected partition field, and obtaining a plurality of
5 incremental data partitions;
performing a partition division on a stock data of the online analytical processing system according to the selected partition field, and obtaining a plurality of stock data partitions; and
integrating data in the incremental data partitions and data in the stock data
10 partitions which have a same partition identity.
2. The method of claim 1, wherein the integrating data in the incremental data partitions and data in the stock data partitions which have a same partition identity comprises:
performing a concatenation operation and/or a union operation on the data in each of
15 the incremental data partitions and the data in the corresponding stock data partitions.
3. The method of claim 2, further comprising:
constructing a data index of the plurality of the stock data partitions; and performing
the concatenation operation and/or the union operation on the data in each of the
incremental data partitions and the data in the corresponding stock data partitions based
20 on the data index of the plurality of the stock data partitions.
4. The method of any one of claims 1 to 3, wherein the partition field comprises anyone or any combinations of a time field, a business field and a custom field.
5. The method of claim 3, wherein the constructing a data index of the plurality of the stock data partitions comprises:
25 constructing the data index in a manner of a binary tree or a hash.
6. The method of claim 3, wherein the performing the concatenation operation and/or the union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions comprises:
30 searching for the data in each of the incremental data partitions according to the data index, in the corresponding stock data partitions; and
performing the concatenation operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is successful.

7. The method of claim 3 or claim 6, wherein the performing the concatenation operation and/or the union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions comprises:

5 searching for the data in each of the incremental data partitions according to the data index, in the corresponding stock data partitions; and

performing the union operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is unsuccessful.

8. The method of any one of claims 1 to 7, after the integrating data in the
10 incremental data partitions and data in the stock data partitions which have a same partition identity, further comprising:

exporting an integrated data to a specific location through transforming and loading in order to perform a data analysis of the online analytical processing system; and/or

15 outputting prompt information, which is used for indicating that the data is updated, to a downstream data processing unit.

9. An apparatus of data processing for online analytical processing, comprising:

a partition divisioning module, configured to perform a partition division on an incremental data of an online analytical processing system according to a selected partition field, and obtain a plurality of incremental data partitions;

20 the partition divisioning module, further configured to perform a partition division on a stock data of the online analytical processing system according to the selected partition field, and obtain a plurality of stock data partitions; and

a data processing module, configured to integrate data in the incremental data partitions and data in the stock data partitions which have a same partition identity.

25 10. The apparatus of claim 9, wherein the data processing module is further configured to

perform a concatenation operation and/or a union operation on the data in each of the incremental data partitions and the data in the corresponding stock data partitions.

11. The apparatus of claim 10, further comprising:

30 a data index constructing module, configured to construct a data index of the plurality of the stock data partitions; and the data processing module performs the concatenation operation and/or the union operation on the data in each of the incremental

data partitions and the data in the corresponding stock data partitions based on the data index of the plurality of the stock data partitions..

12. The apparatus of any one of claims 9 to 11, wherein the partition field comprises anyone or any combinations of a time field, a business field and a custom field.

5 13. The apparatus of claim 11, wherein the data index constructing module is configured specifically to

construct the data index in a manner of a binary tree or a hash.

14. The apparatus of claim 11, wherein the data processing module comprises:

a search submodule, configured to search for the data in each of the incremental data
10 partitions according to the data index, in the corresponding stock data partitions; and

a concatenation operation submodule, configured to perform the concatenation operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is successful.

15 15. The apparatus of claim 14, wherein the data processing module further comprises:

a union operation submodule, configured to perform the union operation on the data in the incremental data partitions and the associated data in the stock data partitions when the search result is unsuccessful.

16. The apparatus of any one of claims 9 to 15, further comprising:

20 a data exporting module, configured to export an integrated data to a specific location through transforming and loading in order to perform a data analysis of the online analytical processing system.

17. The apparatus of any one of claims 9 to 16, further comprising:

25 an information outputting module, configured to output prompt information, which is used for indicating that the data is updated, to a downstream data processing unit after integrating.