

Claims

1. A data processing method, the method comprising:

selecting a part of to-be-processed key-value pairs as hotspot key-value pairs according to a screening rule, wherein a key-value pair comprises a key representing an attribute and a value representing attribute content;

mapping the hotspot key-value pairs to obtain intermediate result key-value pairs corresponding to the hotspot key-value pairs; and

reducing the intermediate result key-value pairs to generate final result key-value pairs for calling; and

when a non-hotspot key-value pair is called, processing the non-hotspot key-value pair by using a reduce function to generate data for calling.

2. The method of claim 1, wherein selecting a part of to-be-processed key-value pairs as hotspot key-value pairs according to a screening rule comprises:

randomly selecting several to-be-processed key-value pairs as hotspot key-value pairs.

3. The method of claim 1, wherein selecting a part of to-be-processed key-value pairs as hotspot key-value pairs according to a screening rule comprises:

randomly selecting a first number of to-be-processed key-value pairs as candidate key-value pairs;

counting a frequency at which each key-value pair among the candidate key-value pairs is called;

arranging the candidate key-value pairs according to the frequencies; and

selecting a second number of key-value pairs from the candidate key-value pairs according to calling frequencies in an order from high to low as hotspot key-value pairs;

wherein the first number is greater than the second number.

4. The method of claim 3, wherein selecting a part of mapping key-value pairs as hotspot key-value pairs according to a screening rule further comprises:

setting a service category condition set of candidate key-value pairs before the step of randomly selecting a first number of to-be-processed key-value pairs as candidate key-value pairs; and

selecting to-be-processed key-value pairs meeting the service category condition set.

5. The method of claim 1, wherein selecting a part of to-be-processed key-value pairs as

hotspot key-value pairs according to a screening rule comprises:

- setting a calling frequency threshold of the hotspot key-value pairs; and
- when the frequency of a key-value pair being called is greater than the calling frequency threshold, setting the key-value pair as a hotspot key-value pair.

6. The method of any one of claims 1 to 5, wherein the method further comprises optimizing the screening rule by using a machine learning model, optionally comprising:

- counting a distribution condition of frequencies of the key-value pairs being called vs. a single attribute by using a clustering algorithm in the machine learning model;

- according to the distribution condition of frequencies of the key-value pairs being called vs. a single attribute, selecting an interval of key values of attribute content at which the frequencies of the key-value pairs being called are not less than a preset frequency threshold; and

- setting the interval of key values of attribute content as a rule condition of the screening rule.

7. The method of claim 6, wherein the method further comprises:

- when a key-value pair of an attribute and a key-value pair of another attribute are called by service systems having a same service code, setting a union set of intervals of key values of attribute content of the key-value pairs of the two attributes as a rule condition of the screening rule.

8. A data processing method, the method comprising:

- mapping to-be-processed key-value pairs to obtain intermediate result key-value pairs corresponding to the to-be-processed key-value pairs, wherein a key-value pair comprises a key representing an attribute and a value representing attribute content;

- selecting a part of the intermediate result key-value pairs as hotspot key-value pairs according to a screening rule; and

- reducing the hotspot key-value pairs to generate final result key-value pairs for calling; and

- when a non-hotspot key-value pair is called, processing the non-hotspot key-value pair by using a reduce function to generate data for calling.

9. A data processing system, comprising:

a screening module configured to select a part of to-be-processed key-value pairs as hotspot key-value pairs according to a screening rule;

a mapping module configured to map the hotspot key-value pairs to obtain intermediate result key-value pairs corresponding to the hotspot key-value pairs, wherein when non-hotspot key-value pairs are called, the mapping module is configured to map the non-hotspot key-value pairs to obtain intermediate result key-value pairs corresponding to the non-hotspot key-value pairs; and

a reducing module configured to reduce the intermediate result key-value pairs to generate final result key-value pairs for calling;

wherein the key-value pair comprises a key representing an attribute and a value representing a numerical value.

10. The system of claim 9, wherein the screening module configured to select a part of to-be-processed key-value pairs as hotspot key-value pairs is specifically configured to:

randomly select several to-be-processed key-value pairs as hotspot key-value pairs.

11. The system of claim 9, wherein the screening module configured to select a part of to-be-processed key-value pairs as hotspot key-value pairs is specifically configured to:

randomly select a first number of to-be-processed key-value pairs as candidate key-value pairs;

count a frequency at which each key-value pair among the candidate key-value pairs is called;

arrange the candidate key-value pairs according to the frequencies; and

select a second number of key-value pairs having maximum calling frequencies from the candidate key-value pairs as hotspot key-value pairs;

wherein the first number is greater than the second number.

12. The system of claim 11, wherein the screening module configured to select a part of mapping key-value pairs as hotspot key-value pairs is further specifically configured to:

set a service category condition set of candidate key-value pairs before the step of randomly selecting a first number of to-be-processed key-value pairs as candidate key-value pairs; and

select to-be-processed key-value pairs meeting the service category condition set.

13. The system of claim 9, wherein the screening module configured to select a part of to-be-processed key-value pairs as hotspot key-value pairs is specifically configured to:

set a calling frequency threshold of the hotspot key-value pairs; and
when the frequency of a key-value pair being called is greater than the calling frequency threshold, set the key-value pair as a hotspot key-value pair.

14. The system of any one of claims 9 to 13, wherein the system further comprises a rule optimization module configured to:

optimize the screening rule by using a machine learning model.

15. The system of claim 14, wherein the rule optimization module is configured to:

count a distribution condition of frequencies of the key-value pairs being called vs. a single attribute by using a clustering algorithm in the machine learning model;

according to the distribution condition of frequencies of the key-value pairs being called vs. a single attribute, select an interval of key values of attribute content at which the frequencies of the key-value pairs being called are not less than a preset frequency threshold; and

set the interval of key values of attribute content as a rule condition of the screening rule.

16. The system of claim 15, wherein the rule optimization module is further configured to:

when a key-value pair of an attribute and a key-value pair of another attribute are called by service systems having a same service code, set a union set of intervals of key values of attribute content of the key-value pairs of the two attributes as a rule condition of the screening rule.

17. A data processing system, comprising:

a mapping module configured to map to-be-processed key-value pairs to obtain intermediate result key-value pairs corresponding to the to-be-processed key-value pairs, wherein when non-hotspot key-value pairs are called, the mapping module is configured to map the non-hotspot key-value pairs to obtain intermediate result key-value pairs corresponding to the non-hotspot key-value pairs;

a screening module configured to select a part of the intermediate result key-value pairs as hotspot key-value pairs according to a screening rule; and

a reducing module configured to reduce the hotspot key-value pairs to generate final result key-value pairs for calling;

wherein the key-value pair comprises a key representing an attribute and a value representing attribute content.