## LISTING OF CHALLENGED CLAIMS

| **Claim 1** | |
|---|---|
| **[1a]** | A method of processing data of a data set over a distributed system, |
| **[1b]** | wherein the data set comprises a plurality of data groups, |
| **[1c]** | the method comprising: partitioning the data of each one of the data groups into a plurality of data partitions that each have a plurality of key-value pairs and |
| **[1d]** | providing each data partition to a selected one of a plurality of mapping functions that are each user-configurable |
| **[1e]** | to independently output a plurality of lists of values for each of a set of keys found in such map function's corresponding data partition to form corresponding intermediate data for that data group and identifiable to that data group, |
| **[1f]** | wherein the data of a first data group has a different schema than the data of a second data group |
| **[1g]** | and the data of the first data group is mapped differently than the data of the second data group so that different lists of values are output for the corresponding different intermediate data, |
| **[1h]** | wherein the different schema and corresponding different intermediate data have a key in common; and |
| **[1i]** | reducing the intermediate data for the data groups to at least one output data group, including processing the intermediate data for each data group in a manner that is defined to correspond to that data group, so as to result in a merging of the corresponding different intermediate data based on the key in common, |
| **[1j]** | wherein the mapping and reducing operations are performed by a distributed system. |

| | |
|---|---|
| **Claim 2** | The method of claim 1, wherein:<br>the at least one output data group is a plurality of output data groups. |
| **Claim 3** | The method of claim 1, wherein:<br>corresponding intermediate data for a data group being identifiable to the data group includes the corresponding intermediate data being stored such that metadata associated with a manner of storage is identifiable to the data group. |
| **Claim 4** | The method of claim 1, wherein:<br>corresponding intermediate data for a data group being identifiable to the data group includes the corresponding intermediate data being stored in association with data that is identifiable to the data group. |
| **Claim 5** | The method of claim 1, wherein:<br>processing the intermediate data for each data group in a manner that is defined to correspond to that data group includes, for each data group, employing an iterator that corresponds to that data group. |
| **Claim 6** | |
| **[6a]** | The method of claim 1, wherein:<br>the intermediate data includes a plurality of grouped sets of key/value pairs;<br>the reducing step is carried out by a plurality of reducers; |
| **[6b]** | the method further comprises:<br>partitioning the intermediate data into a plurality of partitions, each key/value pair of the intermediate data being provided to a separate one of the partitions; and<br>providing the intermediate data of each partition to a separate one of the reducers. |

| | |
|---|---|
| **Claim 7** | |
| **[7a]** | The method of claim 1, wherein:<br>the intermediate data includes a plurality of grouped sets of key/value pairs;<br>the reducing step is carried out by a plurality of reducers; |
| **[7b]** | the method further comprises:<br>partitioning the intermediate data into a plurality of partitions, at least some of the key/value pairs of the intermediate data being provided to more than one of the partitions; and<br>providing the intermediate data of each partition to a separate one of the reducers. |
| **Claim 8** | The method of claim 7, wherein:<br>in the partitioning step, all of the key/value pairs of the intermediate data are provided to all of the partitions. |
| **Claim 9** | |
| **[9a]** | The method of claim 7, wherein:<br>at least some of the reducers include a sort, group-by-key and combine task; |
| **[9b]** | the method further comprises generating and providing metadata for at least some of the mapping, partitioning, combining, grouping and sorting. |
| **Claim 10** | The method of claim 9, wherein: the reducing step includes processing the metadata. |
| **Claim 11** | The method of claim 10, wherein:<br>processing the intermediate data for each data group in a manner that is defined to correspond to that data group includes, for each data group, employing an iterator that corresponds to that data group, wherein the iterator includes providing the associated metadata to the processing of the reducing step. |

| | |
|---|---|
| **Claim 12** | |
| **[12a]** | The method of claim 5, wherein:<br>the intermediate data includes a plurality of grouped sets of key/value pairs;<br>the reducing step is carried out by a plurality of reducers; |
| **[12b]** | for at least one of the reducers, the iterator corresponding to a particular data group, for that reducer, operates according to a different key of a different schema than the iterator corresponding to another particular data group, for that reducer. |
| **Claim 13** | The method of claim 1, wherein:<br>the intermediate data processing step of the reducing step further comprises processing data that is not intermediate data. |
| **Claim 14** | The method of claim 13, wherein:<br>the reducing step is carried out by a plurality of reducers; and<br>the data that is not intermediate data, for a particular reducer, includes data that is associated with another reducer. |
| **Claim 15** | The method of claim 13, wherein:<br>the reducing step is carried out by a plurality of reducers; and<br>the data that is not intermediate data, for a particular reducer, includes data that is associated with that reducer. |
| **Claim 16** | The method of claim 1 wherein the reducing step includes relating the data among the plurality of data groups. |
| **Claim 17** | |
| **[17a]** | A computer system including a plurality of computing devices, the computer system configured to process data of a data set, wherein the data set comprises a plurality of data groups, the computer system comprises at least one processor and memory that are operable to perform the following operations: |

| | |
|---|---|
| **[17b]** | partitioning the data of each one of the data groups into a plurality of data partitions that each have a plurality of key-value pairs and |
| **[17c]** | providing each data partition to a selected one of a plurality of mapping functions that are each user-configurable |
| **[17d]** | to independently output a plurality of lists of values for each of a set of keys found in such map function's corresponding data partition to form corresponding intermediate data for that data group and identifiable to that data group, |
| **[17e]** | wherein the data of a first data group has a different schema than the data of a second data group |
| **[17f]** | and the data of the first data group is mapped differently than the data of the second data group so that different lists of values are output for the corresponding different intermediate data, |
| **[17g]** | wherein the different schema and corresponding different intermediate data have a key in common; and |
| **[17h]** | reduce the intermediate data for the data groups to at least one output data group, including processing the intermediate data for each data group in a manner that is defined to correspond to that data group so as to result in a merging of the corresponding different intermediate data based on the key in common. |
| **Claim 18** | The computer system of claim 17, wherein:<br>the at least one output data group is a plurality of output data groups. |
| **Claim 19** | The computer system of claim 17, wherein:<br>corresponding intermediate data for a data group being identifiable to the data group includes the corresponding intermediate data being stored such that metadata associated with the manner of storage is identifiable to the data group. |

| | |
|---|---|
| **Claim 20** | The computer system of claim 19, wherein:<br>corresponding intermediate data for a data group being identifiable to the data group includes the computer system being configured to store the corresponding intermediate data in association with data that is identifiable to the data group. |
| **Claim 21** | The computer system of claim 17, wherein:<br>the at least one processor and memory being configured to process the intermediate data for each data group in a manner that is defined to correspond to that data group includes, for each data group, the at least one processor and memory being configured to employ an iterator that corresponds to that data group. |
| **Claim 22** | |
| **[22a]** | The computer system of claim 21, wherein:<br>the intermediate data includes a plurality of grouped sets of key/value pairs;<br>the at least one processor and memory are further operable to reduce the intermediate data via a plurality of reducers; |
| **[22b]** | for at least one of the reducers, the iterator corresponding to a particular data group, for that reducer, is configured to operate according to a different key of a different schema than the iterator corresponding to another particular data group, for that reducer. |
| **Claim 23** | |
| **[23a]** | The computer system of claim 17, wherein:<br>the intermediate data includes a plurality of grouped sets of key/value pairs;<br>the at least one processor and memory are further operable to reduce the intermediate data via a plurality of reducers; and<br>the at least one processor and memory are further operable to: |

| | |
|---|---|
| **[23b]** | partition the intermediate data into a plurality of partitions, to provide each key/value pair of the intermediate data to a separate one of the partitions; and<br>provide the intermediate data of each partition to a separate one of the reducers. |
| **Claim 24** | |
| **[24a]** | The computer system of claim 17, wherein:<br>the intermediate data includes a plurality of grouped sets of key/value pairs;<br>the at least one processor and memory are further operable to reduce the intermediate data via a plurality of reducers; and<br>the at least one processor and memory are further operable to: |
| **[24b]** | partition the intermediate data into a plurality of partitions, to provide at least some of the key/value pairs of the intermediate data to more than one of the partitions; and<br>provide the intermediate data of each partition to a separate one of the reducers. |
| **Claim 25** | The computer system of claim 24, wherein:<br>the at least one processor and memory are further operable to, in partitioning, to provide all of the key/value pairs of the intermediate data to all of the partitions. |
| **Claim 26** | |
| **[26a]** | The computer system of claim 24, wherein:<br>at least some of the reducers include a sort, group-by-key and combine task;<br>the at least one processor and memory are further operable to |
| **[26b]** | generate and provide metadata for at least some of the mapping, partitioning, combining, grouping and sorting. |

| | |
|---|---|
| **Claim 27** | The computer system of claim 26, wherein:<br>the reducing includes processing the metadata. |
| **Claim 28** | The computer system of claim 27, wherein:<br>the at least one processor and memory being further operable to process the intermediate data for each data group in a manner that is defined to correspond to that data group includes, for each data group, the at least one processor and memory being further operable to employ an iterator that corresponds to that data group, wherein the iterator is configured to provide the associated meta data to the reducing. |
| **Claim 29** | The computer system of claim 17, wherein:<br>the intermediate data processing of the reducing further comprises processing data that is not intermediate data. |
| **Claim 30** | The computer system of claim 29, wherein:<br>the at least one processor and memory are further operable to reduce the intermediate data via a plurality of reducers; and<br>the data that is not intermediate data, for a particular reducer, includes data that is associated with another reducer. |
| **Claim 31** | The computer system of claim 29, wherein:<br>the at least one processor and memory are further operable to reduce the intermediate data via a plurality of reducers; and<br>the data that is not intermediate data, for a particular reducer, includes data that is associated with that reducer. |
| **Claim 32** | The computer system of claim 17, wherein the at least one processor and memory are further operable to reduce by relating the data among the plurality of data groups. |
| **Claim 33** | |
| **[33a]** | A map-reduce method of processing data from a plurality of groups having different schema over a computer system, the method comprising: |

| | |
|---|---|
| **[33b]** | for a first data set having a plurality of first key-value pairs, wherein such first data set belongs to a first data group and the first key-value pairs have a first schema, |
| **[33c]** | partitioning the first data set into a plurality of data partitions and |
| **[33d]** | providing each data partition to a selected one of a plurality of mapping functions that are each user-configurable |
| **[33e]** | to independently output a plurality of lists of values for each of a set of keys found in such map function's corresponding data partition to form a first intermediate data set having a first set of resulting key-value pairs; |
| **[33f]** | for a second data set having a plurality of second key-value pairs, wherein such second data set belongs to a second data group and the second key-value pairs have a second schema, |
| **[33g]** | partitioning the first data set into a plurality of data partitions and |
| **[33h]** | providing each data partition to a selected one of a plurality of mapping functions to form a second intermediate data set having a second set of resulting key-value pairs differing from the first set of resulting key-value pairs, |
| **[33i]** | wherein the first schema differs from the second schema, |
| **[33j]** | wherein the first and second schema and the first and second set of resulting key-value pairs have a key in common; and |
| **[33k]** | reducing the first and the second intermediate data set together so as to form an output data set, wherein the reducing is accomplished by iterating on at least one key from the first and second intermediate data set, and |
| **[33l]** | the output data set has a different schema than the first and second schema, |

| | |
|---|---|
| **[33m]** | wherein the mapping and reducing operations are performed by a distributed system. |
| **Claim 34** | The map-reduce method of claim 33, wherein the reducing is accomplished by iterating on at least one key from each of the first and second intermediate data set so that the output data set is a merging of a portion of the first and second intermediate data set. |
| **Claim 35** | The map-reduce method of claim 33, wherein the reducing is accomplished by a plurality of reducers, the method further comprises:<br>partitioning the first or second intermediate data set into a plurality of partitions, at least some of the key-value pairs of the first or second intermediate data being provided to more than one of the partitions; and<br>providing the partitioned intermediate data set of each partition to a separate one of the reducers. |
| **Claim 36** | The map-reduce method of claim 35, wherein all of the key-value pairs of the first and second intermediate data set are provided to all of the reducers. |
| **Claim 37** | The map-reduce method of claim 35, wherein at least some of the reducers include a sort, group-by-key and combine task, the method further comprises generating and providing metadata for at least some of the mapping, partitioning, combining, grouping and sorting. |
| **Claim 38** | The map-reduce method of claim 37, wherein the reducing includes processing the metadata. |
| **Claim 39** | The map-reduce method of claim 38, wherein iterating includes providing the associated metadata to the processing of the reducing step. |
| **Claim 40** | |
| **[40a]** | A computer system including a plurality of computing devices, the computer system configured to process data of a data set, |

| | |
|---|---|
| | wherein the data set comprises a plurality of data groups, the computer system comprises at least one processor and memory that are operable to perform the following operations: |
| **[40b]** | for a first data set having a plurality of first key-value pairs, wherein such first data set belongs to a first data group and the first key-value pairs have a first schema, |
| **[40c]** | partitioning the first data set into a plurality of data partitions and |
| **[40d]** | providing each data partition to a selected one of a plurality of mapping functions that are each user-configurable |
| **[40e]** | to independently output a plurality of lists of values for each of a set of keys found in such map function's corresponding data partition to form a first intermediate data set having a first set of resulting key value pairs; |
| **[40f]** | for a second data set having a plurality of second key-value pairs, wherein such second data set belongs to a second data group and the second key-value pairs have a second schema, |
| **[40g]** | partitioning the first data set into a plurality of data partitions and |
| **[40h]** | providing each data partition to a selected one of a plurality of mapping functions to form a second intermediate data set having a second set of resulting key-value pairs differing from the first set of resulting key-value pairs, |
| **[40i]** | wherein the first schema differs from the second schema; and |
| **[40j]** | reducing the first and the second intermediate data set together so as to form an output data set, wherein the reducing is accomplished by iterating on at least one key from the first and second intermediate data set |
| **[40k]** | and the output data set has a different schema than the first and second schema, |

| | |
|---|---|
| **[40l]** | wherein the first and second schema and the first and second set of resulting key-value pairs have a key in common. |
| **Claim 41** | |
| **Claim 42** | The computer system of claim 40, wherein the reducing is accomplished by a plurality of reducers, the d the at least one processor and memory further operable for:<br>partitioning the first or second intermediate data set into a plurality of partitions, at least some of the key-value pairs of the first or second intermediate data being provided to more than one of the partitions; and<br>providing the partitioned intermediate data set of each partition to a separate one of the reducers. |
| **Claim 43** | The computer system of claim 42, wherein all of the key-value pairs of the first and second intermediate data set are provided to all of the reducers. |
| **Claim 44** | The computer system of claim 42, wherein at least some of the reducers include a sort, group-by-key and combine task, the d the at least one processor and memory are further operable for:<br>generating and providing metadata for at least some of the mapping, partitioning, combining, grouping and sorting. |
| **Claim 45** | The computer system of claim 44, wherein the reducing includes processing the metadata. |
| **Claim 46** | The computer system of claim 45, wherein iterating includes providing the associated metadata to the processing of the reducing step. |