

(51) International Patent Classification:
G06F 9/50 (2006.01)

(21) International Application Number:

PCT/EP2016/000713

(22) International Filing Date:

02 May 2016 (02.05.2016)

(25) Filing Language:

English

(26) Publication Language:

English

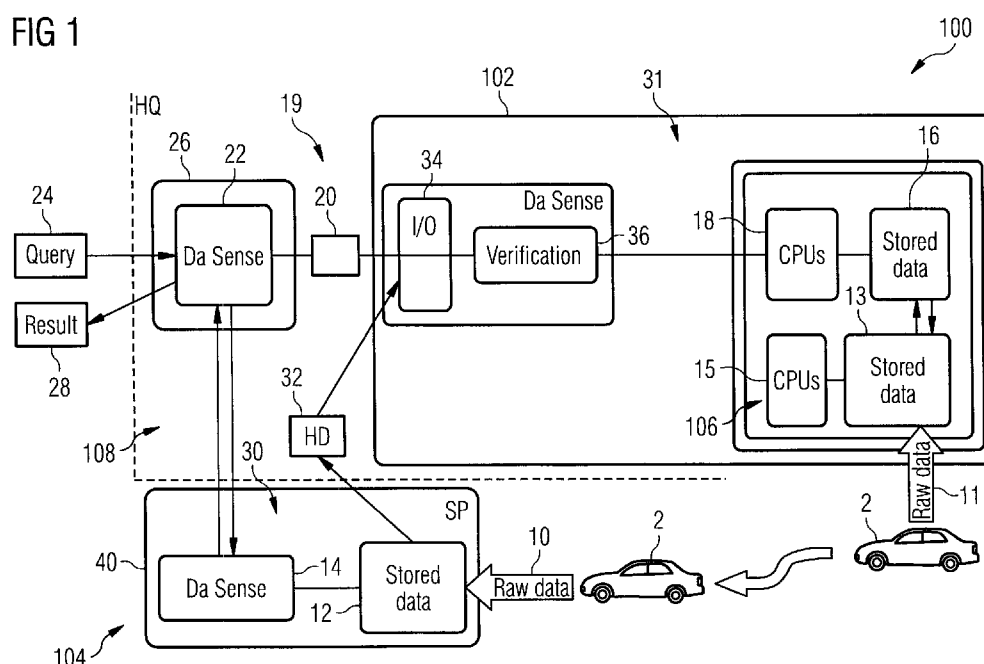
(71) Applicant: **NORCOM INFORMATION TECHNOLOGY AG** [DE/DE]; Gabelsbergerstrasse 4, 80333 München (DE).(72) Inventor: **ABTHOFF, Tobias**; c/o NorCom Information Technology AG, Gabelsbergerstrasse 4, 80333 München (DE).(74) Agent: **KRAMER BARSKE SCHMIDTCHEN**; Patentanwälte PartG mbB, Landsberger Strasse 300, 80687 München (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

(54) Title: DISTRIBUTED DATA ANALYSIS SYSTEM AND METHOD

FIG 1



(57) **Abstract:** The present invention relates to a distributed data analysis system (100) for analyzing large amounts of collected measurement data, for example, data that is accumulated during test drives of vehicles (2) in the field of automotive engineering. When a query for an analysis to be performed on the collected measurement data is received, an analysis device (26) determines which of a plurality of different storage devices (12, 16) located at geographically different locations include measurement data relevant to the query, and performs the analysis on the measurement data stored on the relevant storage devices. The partial results of the analysis are combined and reported back to a user. Data is transferred between the storage devices (12, 16) based on, for example, a remaining storage capacity of the same or a priority of the measurement data.



TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

DISTRIBUTED DATA ANALYSIS SYSTEM AND METHOD

Technical Field

The present invention generally relates to a distributed data analysis system and method, in particular, to a distributed data analysis system and method for use in analyzing measurement data in the field of automotive engineering.

Background

During development of a new car model, there are several development phases. In particular, during a prototype phase, many prototype cars are driven large distances of up to around 150.000 km within a few weeks in order to evaluate a long-term behavior and wear of the components of the car. Usually, the cars are driven at several locations around the globe to probe the influence of, for example, climate, fuel composition and the like. During such test drives, a large amount of measurement data is accumulated and stored. For example, several Gigabyte (GB) of data are collected per hour of driving. This is a result of the large number of different measurement signals that are processed (at present, around 1.600, but up to 20.000 or more in the future). In some cases, a so-called full bus trace is recorded, i.e., all signals sent through the bus system of the car are stored in temporal order. Currently, this amounts to 1 GB per hour for a CAN trace, and even more for other bus systems like FlexRay or Ethernet.

The collected data is offloaded at regular intervals at geographically distributed service points, mostly maintained by external service providers. These providers may be automotive companies that do not have a sophisticated IT infrastructure and may only have an average or poor internet connection. Therefore, it may take a long time to upload the data to the headquarters (a central office or the like) of the car manufacturer.

Similarly, in the field of autonomous driving, an even larger amount of sensory data is accumulated during testing of a self-driving car. For example, the rate of incoming sensory data may be around 1-2 GB per second. Currently, algorithms that have to be developed and tested are applied to individual sets of test drive data by individual engineers, where the algorithm test runs are performed on powerful workstations and the data is stored on an external hard drive. Clearly, this limits the amount of data that can be analyzed. Further, whenever an algorithm is to be tested with different parameters, this can only be achieved after a previous test run has finished.

Therefore, it is desirable to provide a system and method that allow for a flexible and prompt analysis of large amounts of measurement data, in particular, in the field of automotive engineering and autonomous driving, but not limited to the same

Summary of the Invention

In a first aspect of the present invention, a distributed data analysis system for analyzing collected measurement data comprises a data input device configured to receive measurement data, a first storage device associated with the data input device and configured to store the measurement data input via the data input unit, and a first computing device associated with the first storage device. A second storage device is configured to store measurement data previously stored on the first storage device. A second computing device is associated with the second storage device, and a data distribution system is configured to distribute the measurement data between the first storage device and the second storage device based on at least one predetermined criterion. A data management device is configured to store a location of the measurement data and to update the stored location based on the distribution by the data distribution device. A query input device is configured to receive a query for an analysis to be performed on the collected measurement data. An analysis device is configured to perform the

analysis on the measurement data stored on the first storage device and the second storage device in parallel, preferably simultaneously, by the first computing device and the second computing device, respectively, based on the location of the collected measurement data stored by the data management device. A reporting device is configured to report a result of the analysis of the collected measurement data.

Such a distributed data analysis system offers the advantage that, in the context of automotive engineering, the data that is offloaded by a test vehicle at an arbitrary location around the globe and stored on the first storage device does not need to be transferred to the headquarters before an analysis can be performed on said measurement data. Instead, when a query is received and an analysis is to be performed, part of the analysis is performed at the location of the first storage device, i.e., on the local measurement data offloaded by the test vehicle. This is possible because the data management device continually updates the location of all available measurement data, such that an analysis to be performed can be distributed to both the first computing device associated with the first storage device at the external location and the second computing device associated with the second storage device located at the headquarters and containing the bulk of the collected measurement data. It will be readily apparent that generally a plurality of first storage devices at geographically different locations, each being associated with a corresponding data input device and a corresponding first computing device, will be provided, and the analysis may be distributed to all the service points at the geographically different locations.

In a further aspect, a method for distributed data analysis of collected measurement data stored on a first storage device and a second storage device comprises receiving measurement data, storing a location of the received measurement data, and distributing the measurement data between the first storage device and the second storage device based on at least one predetermined criterion. The method further comprises updating the location of the measurement

data based on the distribution, receiving a query for an analysis to be performed on the collected measurement data, and performing the analysis on the measurement data stored on the first storage device and the second storage device by a first computing device associated with the first storage device and a second computing device associated with the second storage device, respectively, in accordance with the stored location of the measurement data. In a further step, a result of the analysis is reported.

Using the system and method disclosed herein, the data does not need to be transferred, i.e., uploaded to the headquarters via the internet. Further, it is also not necessary to send the data via physical mail to the headquarters before it can be analyzed. Instead, the query or analysis software, i.e., the algorithm or program for performing the analysis, is sent to the local service points. This process is transparent to a user of the data analysis system, in other words, the user does not know which part of the measurement data is available at the headquarters, and which part is only available at the local service points. After the analysis has been completed, the result is displayed to the user, irrespective of where the partial analyses have been performed.

Within the context of autonomous driving, the distributed data analysis system of the present invention allows, among other things, solving the storage problem associated with the enormous amount of data that is generated during a test drive. Further, much more computing power is made available to an engineer for performing algorithm test runs. This is achieved by providing computing devices having different ratios of computing power (for example, number and quality of CPU cores and main memory) to storage capacity (number and size of hard disks). For example, the first computing device may have more computing power than the second computing device, and the second storage device may be co-located with the first storage device and be considerably larger than the same. The data distribution system may be configured to classify the measurement data resulting from different test drives into data having different priorities, and to

transfer data having a lower priority to the second storage device. In this manner, for example, most recent data may be stored on the first storage device associated with the computing device having a higher computing power, as it represents “hot” data that is most likely accessed in the near future during testing of algorithms. The data on the second storage device may be data having a lower priority, i.e., data that is less likely to be accessed in the immediate future, or data that is accessed less frequently, for example, older than the “hot” data (e.g., “warm” data).

Also in the context of autonomous driving, a distributed data analysis may be performed by the computing device associated with the storage device having the high priority data and the computing device associated with the storage device having the low priority data. Further, data that is outdated, but has to be stored for various reasons, for example, in case of an analysis to be performed at a later time, can be transferred to a further storage device serving as an object store that has practically no computing power. However, meta data is generated for the outdated data (referred to as “frozen” data), and in case it is necessary to analyze part of the “frozen” data, said data can be transferred to the other storage devices associated with the respective computing devices on the basis of the meta data.

In this respect, a classification of the data to be analyzed may be based on access times, creation dates, other meta data or other content-related criteria. Advantageously, the different computing devices and storage devices are part of a cluster computing framework such as a Hadoop cluster. Different server groups may be defined in the cluster by the ratio of computing power to storage capacity and may serve to store, for example, the “hot” data, the “warm” data and the “frozen” data, respectively.

In case of an analysis such as a parameter test or an optimization to be performed on the measurement data, the analysis may be performed, for example, on both the “hot” data and the “warm” data by the respective compute nodes in

the cluster. In particular, a parallel analysis/simulation using different parameters may be performed by running several instances of the algorithm on several cluster compute cores.

In another aspect, a computer program product comprises computer-executable instructions that, when executed on a computer system, cause the computer system to execute the steps of acquiring a location of measurement data stored on a plurality of storage devices provided at geographically different locations, receiving a query for an analysis to be performed on the measurement data, evaluating the query to determine on which of the plurality of storage devices data relevant to the query is stored, generating analysis code based on the query, forwarding the generated analysis code to computing devices associated with the relevant storage devices, respectively, receiving a partial analysis result from each computing device, combining the partial analysis results, and reporting the combined analysis result.

Brief Description of the Drawings

In the following, embodiments of the present invention are described in detail with reference to the attached drawings, in which:

Fig. 1 shows a schematic overview of a distributed data analysis system in accordance with an embodiment of the present invention;

Fig. 2 shows a block diagram illustrating a processing performed in accordance with the embodiment of the present invention; and

Fig. 3 shows a schematic overview of a distributed data analysis system in accordance with another embodiment of the present invention.

Detailed Description

Fig. 1 shows a schematic overview of a distributed data analysis system 100 for analyzing collected measurement data. In particular, distributed data analysis system 100 is a system for use in the analysis of measurement data

generated by one or more vehicles 2 during test drives associated with the development of a new car model or after introduction of a new car model into the market.

As shown in Fig. 1, distributed data analysis system 100 comprises a central computing system 102 provided, for example, at the headquarters (HQ) of the car manufacturer, and a plurality of decentralized service points (SP) or service providers (i.e., computing systems) 104, only one of which is shown in Fig. 1. The plurality of decentralized service points are located at geographically different locations around the globe and visited by one or more test vehicles during test drives of the same. When a test vehicle 2 stops at one of service points 104, it offloads the measurement data (raw data) acquired during a test drive at the service point. It should be noted that Fig. 1 shows a further service point 106 co-located with main computing system 102. In other embodiments, however, service point 106 may be omitted.

Each service point 104 includes a data input device 10 for receiving the measurement data generated during the test drive. The amount of data typically will be several GB or several tens of GB. Each service point 104 includes a storage device 12, for example, one or more hard disks for storing the measurement data input via data input device 10. It will be appreciated that any appropriate storage device having a storage capacity that is large enough to store at least several hundreds of GB of data may be used. Service point 104 further includes a computing device 14 associated with storage device 12. Computing device 14 may be any known computing device including one or more CPUs, a main memory, and the like, and being configured to execute program code. In some embodiments, computing device 14 and storage device 12 may be part of a computing cluster 30, for example, a Hadoop cluster. This has the advantage that it also provides basic data security via data replication on several hard disks in the cluster.

Storage device 12 and computing device 14 may be contained in a housing 40 forming a standalone unit and including data input device 10. It will be readily appreciated that an appropriate operating system is provided for storage device 12 and computing device 14 to facilitate processing of the data on storage device 12 by computing device 14, as well as processing of the raw data input via data input device 10 and communication between computing device 14 and the outside, for example, via a data link 108 such as the internet. Service point 106 also includes a storage device 13 and an associated computing device 15 configured in a manner that is similar to the configuration of storage device 12 and computing device 14.

Main computing system 102 includes a main storage device 16 and an associated main computing device 18. Storage device 16 has a much larger storage capacity than, for example, storage devices 12, 13. As such, storage device 16 serves as a “data lake” storing the bulk of the available measurement data collected by one or more vehicles. Also in this case, storage devices 13, 16 and computing devices 15, 18 may be part of a large computing cluster 31, for example, a Hadoop cluster.

Generally, the measurement data offloaded at service point 106 can be easily transferred to main storage device 16 due to service point 106 being co-located with main computing system 102. However, it will be readily apparent that this may not always be possible for service point 104, which may be located anywhere around the globe, for example, in areas with poor internet connection. Therefore, the data offloaded at service point 104 may not be available for analysis on main computing system 102 in a timely manner. Therefore, in the present embodiment, part of the analysis of the measurement data is performed at local service point 104, in combination with an analysis that is performed in parallel on the measurement data available at the headquarters. This will be described in more detail below.

When raw data is input via data input device 10 at service point 104, the raw data may be converted into a data format suitable for a particular data analysis application, for example, one that supports distributed processing of data on a plurality of computing devices. For example, so-called Big Data applications such as Spark may be used for the analysis. To this end, for example, the proprietary software DaSense by NorCom Information Technology AG may run on computing device 14 of service point 104 (and also on main computing system 102). In particular, the raw data obtained from the test vehicle is transferred to a particular region of the memory device 12 of cluster 30 (a so-called “landing zone”) by the software. This memory region is scanned for new data on a regular basis, and when new data is found, an ingest process is triggered automatically by the software. A specific ingest (read) process, e.g., a specific data conversion process is selected depending on the format and the type of raw data input via data input device 10. The format and type of data may either be input by an operator or may be automatically detected by computing device 14. Using the software, the proprietary and often complex automotive data formats are parsed and transformed into data structures that are suitable for the distributed analysis that is to be performed. After conversion, the data is moved to its long-term storage location on memory device 12. In some embodiments, the raw data may also be stored on memory device 12. In other embodiments, no conversion may be performed, and the raw data may be stored as is, such that the analysis software must be capable of performing a query/analysis on the raw data.

During the above-described ingest process, meta data is generated that indicates, among others, the type of data that has been newly stored and the location of the same. For example, the meta data may include an identification of the vehicle that has offloaded the data, a date and time, the contents and/or format of the data, an identification code of service point 104, a memory location of the data on storage device 12, and the like. Further, during the ingest process, an initial analysis may be performed, e.g. plausibility and quality checks, evaluation

of a range of statistics or pre-defined criteria and the like. In case of problems, for example, poor data quality, problematic or irregular data, failed data upload or corrupted data due to a damaged hard disk in the vehicle, alerts are generated that are uploaded to main computing system 102 via data link 108. Once the measurement data is ingested (which may take several minutes or hours), their existence is reported back to the main computing system 102 at the headquarters.

To this end, distributed data analysis system 100 includes a data management device 22 that is in communication with computing device 14 of each service point 104, as well as with main computing system 102. For example, data management device 22 may include a web server running the DaSense software. Data management device 22 is configured to store a location of the measurement data that is offloaded at each service point 104 and that is stored on main storage device 16, for example, in an appropriate database. To this end, data management device 22 receives the meta data generated by service point 104 during the data ingest process and forwards the same to main computing system 102. In this manner, the location of all the data that is available for performing a particular analysis is stored, for example, on main computing system 102.

Distributed data analysis system 100 further includes a query input device 24, an analysis device 26 and a reporting device 28.

Query input device 24 may be a conventional personal computer or workstation used by an engineer formulating the query. For example, the query may be formulated in the DaSense domain specific language (DSL) or any other appropriate programming language. Alternatively or additionally, a graphical user interface may be provided for inputting the query. Typical queries might be, for example, a query for finding all measurements that were taken during a specific period of time with a particular type of engine, an electronic control unit (ECU) running on a particular software version and with a temperature difference before and after a diesel particle filter (DPF) that is larger than 200 °C. Another typical query might be a query for identifying measurements that have similar

patterns to a specific measurement in a specified range of measurement channels and fulfilling certain meta data requirements (date, time, etc.). In another exemplary query, it may be desired to combine all available measurements for a specific car and return the durations and the periodicity of the DPF regeneration events, as well as some properties of the oil in the car during the regenerations (it should be noted that oil properties are generally not measured in the car, but measured regularly by sending oil probes to a lab, and hence are stored separately on a separate storage device, which may form part of distributed data analysis system 100).

In some embodiments, the query may be for performing a full trace analysis for one or more specific cars. In this case, an appropriate program or algorithm for performing the query (i.e. the analysis) on the data stored, for example, on storage device 12 is downloaded to computing device 14 and executed by the same. In particular, in a first stage, the (relevant) state history of the bus is inferred by the software, and the tests are performed in a parallel fashion on the cluster 30 in a second stage. Test cases may be cases where, for example, the time difference between subsequent network messages from each ECU fall in a predefined range during the bus wake state, no non-network management messages are to be sent to the bus during the bus sleep state, at the beginning of the bus wake state, the network messages have to have a certain predefined payload specifying the reasons for wake-up, and the like.

The query input via query input device 24 is sent to analysis device 26. Analysis device 26 may also include a web server running the DaSense software or any other server running an appropriate software. Analysis device 26 evaluates the meta data stored, for example, on main computing system 102 to determine the location of the measurement data in order to identify where the relevant measurement data is stored, and then simultaneously performs the analysis on the measurement data stored, for example, on main storage device 16 and remote

storage device 12 and being relevant for the query. This will be described in more detail in the following with reference to Fig. 2.

Fig. 2 shows a block diagram describing a processing that is performed by distributed data analysis system 100 when a query is received. In particular, Fig. 2 shows the detailed configuration of data analysis device 26 and an associated data distribution system 19.

As shown in Fig. 2, data analysis device 26 includes a query optimizer 50 and an asynchronous distributed execution coordinator 64. Query optimizer 50 is configured to receive the query from query input device 24 and to initiate the distributed analysis based on the received query. To this end, query optimizer 50 is in communication with data management device 22 and configured to retrieve, for example, the size and location of the available measurement data. Further, query optimizer 50 is configured to generate a distributed execution graph 60 and a pushdown plan 62 comprising computer-executable instruction that configure the distributed analysis that is to be performed. In order to determine the most efficient manner in which the analysis is to be performed, query optimizer 50 is connected to an algebraic library 52, a cost prediction module 54, an execution engine capabilities module 56 and a bandwidth graph module 58.

Algebraic library 52 includes a set of rules for modifying directed execution graphs governing the processing of a query. Generally, such directed execution graphs include a plurality of operators to be applied to data that is to be analyzed, for example, in a given order. The set of rules included in algebraic library 52 specifies how said operators may be manipulated, for example, commutated with each other. In this manner, an operator algebra is defined. For example, in a simple query, two sets of data may need to be processed, e.g., a first table including meta data specifying, for example, all available test vehicles and their respective configurations and a second, generally much larger table including measurement data. Exemplary operators may be a scan operator applied to the respective data sets for determining which data are relevant to the query.

Then, a join operator may operate on the scanned data sets and join the same, resulting in a smaller data set that only includes the relevant data. In a subsequent set, for example, a filter operation may be applied to further reduce the amount of data based on predetermined criteria (for example, only data that has been generated within a predetermined period of time or the like). Finally, the filtered data may be grouped and sorted to obtain the query result.

In the above-described scenario with the plurality of test vehicles performing test drives around the globe, however, the different data sets may only be available at different locations. For example, if the measurement data are only available at a remote service point such as service point 104, a distributed execution graph needs to be generated for performing the scan operation of the measurement data at the remote service point. Further, the scanned measurement data needs to be transferred to the location where the other data are located in order to perform the join operation. This results in a large amount of data that needs to be transferred, resulting in considerable delays, for example, when the data upload rates are limited.

According to the present invention, the processing may be made more efficient by modifying the distributed execution graph, for example, by a commutation of operators, e.g., moving operators from one location to another in accordance with the predefined operator algebra. For example, in order to reduce the amount of data than needs to be uploaded, the filter operation may be performed at the remote location after the scan of the measurement data. That is to say, the filter operation is performed not after the join operation, but before the same. The filtered measurement data are then uploaded to the location of the other data, for example, the table including the meta data, and joined with the same.

In a further step, the join operation may also be performed at the remote location, for example, at each service point. This requires that data has to be transferred from the central location to the remote location in order to perform the

join operation. After the join operation, the joined data is uploaded to the central location. Here, it may be more efficient to first download some data from the central server to the service point, and to only upload the joined data, than to perform the join operation on the central server.

In some applications, the filter operation may be duplicated, i.e., performed both at the remote location and the central location on the respective scanned data sets. This may further reduce the amount of data that has to be transferred from the central location to the remote location.

In order to determine the optimum configuration of the distributed execution graph, query optimizer 50 is in communication with cost prediction module 54, execution engine capabilities module 56 and bandwidth graph module 58. Cost prediction module 54 is configured to include a set of rules, on the basis of which the execution costs of a given query, e.g., the required execution time, the required CPU power and required bandwidth can be predicted. Execution engine capabilities module 56 includes information on the speed and capabilities of the distributed execution engines (for example, computing device 14 at service point 104), wherein the speed is used for the cost prediction and the capabilities are used for determining which operations can be performed by the engine. Each execution engine is generally capable of performing certain operations on the respective data, accepting data and delivering data.

Bandwidth graph module 58 includes information on the usable bandwidth for the connection to the respective execution engines, where separate values may be stored for upload and download. This information is also used for the cost prediction by query optimizer 50.

Previously described data management device 22 includes information on the size and location of the available measurement data, where the data size is used for both query result size estimation and query CPU cost estimation. The data location is used for, among other things, the bandwidth cost estimation.

Query optimizer 50 optimizes the analysis based on the query by applying the algebraic rules specified in algebraic library 52 and performs an appropriate pushdown planning, i.e., an appropriate distribution of individual jobs to the available execution engines. The optimization is performed according to the predicted costs that are determined in the above-described manner. The distributed execution graphs generated in this manner are then forwarded to asynchronous distributed execution coordinator 64 that pushes the respective jobs to the distributed execution engines inbound queues, collects results from the distributed execution engines, and the like.

As outlined above, a pushdown plan module 62 may transfer some data before executing certain distributed execution graphs. In particular, each pushdown should be formulated as a function with associated costs such that it can be manipulated by query optimizer 50.

Data distribution system 19 may include a data distribution device 20 that listens to the pushdown requests of pushdown plan module 62 to find optimal data placements by moving or replicating some data. For example, some data that is frequently downloaded to remote locations during processing could be permanently replicated on the remote system, for example, at service point 104 shown in Fig. 1, thus acting as a local cache. In this respect, however, it is important to note that the distribution by data distribution device 20 is also governed by a set of data governance rules 66, which may be a set of manual rules describing, for example, a desired maximum local data size on each service point, a desired maximum local data age at each service point, data that is not to be moved/replicated to a service point, for example, due to security concerns, data that has to be replicated in a mandatory fashion, for example, to provide local fail-safe capabilities, etc. Based on the data governance rules 66 and pushdown plan 62, data management device 20 computes desired data locations and forwards the computed desired data locations to a data movement planner 68. Data movement planner 68 generates an execution plan for moving data from one

location to another location. For example, data movement planner 68 may determine that data needs to be moved from service point 104 to central computing system 102, e.g., due to the maximum local data size on storage device 12 being exceeded. Data movement planner is configured to choose from moving the data in an online or an offline manner, i.e., by transferring the same via data link 108 or via physical mail, for example, using portable hard drives 32 sent, for example, via DHL or a similar courier service. The corresponding determinations are forwarded to an online move queue 70 and an offline move queue 72. As indicated by the arrow in Fig. 2, the locations of the moved data are continuously updated and stored by data management device 22 to be used in subsequent queries.

After the analysis has been performed by the respective computing devices, for example, by performing the distributed analysis described above, the results of the partial analyses are collected by analysis device 26 (more precisely, asynchronous distributed execution coordinator 64) and combined to create a query result. The query result is then forwarded to reporting device 28 that is configured to report the result of the analysis of the collected measurement data, for example, to the user that has input the query.

It will be readily appreciated that, in other embodiments, specific queries may be automatically generated by distributed data analysis system 100, for example, standard queries for certain car behaviors, car locations, data types and the like, which may be generated on a regular basis, and the results of the queries may be stored for future reference without immediately being reported to a user. For example, predefined report may be generated after lapse of a predetermined time period, for example, on a weekly basis, and the available data may be retrieved by engineers in a web client or as a PDF document at a later time.

Generally, in the above processing, first a meta data query is evaluated. In other words, it is determined to which clusters the query has to be forwarded because they have eligible data. For example, it may be determined that the query

has to be forwarded to main computing device 18 because main storage device 16 holds eligible data, and to several service points 104 around the globe, which include storage devices 12 that also hold eligible data. By first evaluating the meta data query in this manner, the amount of eligible data that has to be evaluated can be reduced dramatically. Then, for example, an Apache Spark job or another appropriate program is started on each cluster in order to conduct the analysis on the respective data. The result is transferred back to analysis device 26 via data link 108, and analysis device 26 combines the partial results and generates the final result as described above.

It will be obvious that each service point 104 cannot provide unlimited storage space on storage device 12. Therefore, to ensure long-term data security, the geographically distributed data is sent to the headquarters by physical mail, for example, DHL, to be transferred to main computing system 102 (in particular, main storage device 16). To this end, distributed data analysis system 100 includes data distribution system 19 configured to transfer the measurement data from the respective storage devices 12 at the geographically different locations to the main storage device 16 at the headquarters, for example, based on at least one pre-determined criterion. As mentioned above, the at least one criterion may include at least one of a remaining capacity of one of the storage devices 12, lapse of a pre-determined time interval, for example one week, and a specific type of the measurement data stored on the respective storage devices 12, for example, measurement data having a high priority, or a likelihood of being frequently used in future analyses. For example, data distribution system 19 may include data distribution device 20 configured to initiate a backup of at least part of the measurement data stored on first storage device 12 to a portable storage device 32 at the location of first storage device 12 when the at least one criterion is met. Although data distribution device 20 is shown as being provided separately, it will be readily appreciated that data distribution device 20 may also be provided as part of main computing system 102 or other components of

distributed analysis system 100, for example, analysis device 26. Further, in other embodiments, an operator may determine that a data backup is necessary and initiate the same in order to send the data to the headquarters.

When the data is stored on the portable storage device 32, meta data indicating the origin of the data, i.e., the identity of the service point 104, the date of creation of the backup, the type and location of the data on storage device 12, and the like may be generated and associated with portable storage device 32. For example, a bar code or other identifier may be created and printed to be attached to portable storage device 32. Alternatively, said meta data could also be stored in a designated area of portable storage device 32. Preferably, the data stored on portable storage device 32 is encrypted to ensure data security.

After portable storage device 32 has been physically transported to the headquarters, it is connected to a backup data input device 34 provided at the location of main computer system 102, i.e., main storage device 16, and read by backup data input device 34. In order to transfer the backup data to second storage device 16 in this process, a data verification device 36 may verify that the data that has been transferred to the second storage device 16 is identical to the measurement data previously stored on storage device 12, for example, by comparing a hash sum of the data, a list of files, file sizes, checksums, etc., that may be part of the meta data generated during backup, and may initiate deletion of the measurement data previously stored on the storage device 12 after successful verification. To this end, data verification device 36 may communicate with data management device 22 such that data management device 22 updates the location of the verified measurement data for future access.

Turning to Fig. 3, another embodiment of a distributed data analysis system 200 in accordance with the present invention is shown.

As shown in Fig. 3, distributed data analysis system 200 comprises a main computing system 202, a data management device 122, a data distribution system 119 including a data distribution device 120, a query input device 124, a

reporting device 128, an analysis device 126, and a main computing cluster 131. Distributed data analysis system 200 is configured similarly to distributed data analysis system 100 such that only the differences will be described in detail in the following.

Data analysis system 200 is suitable for use in an autonomous driving application. In such an application, a large number of algorithms has to be developed to interpret incoming sensory data from, e.g., cameras, radar or lidar systems or the like in order to maintain an accurate representation of the vehicle state and its environment. In order to test such algorithms, the sensory data, which normally has to be analyzed in real time, is recorded, such that new versions of an algorithm can be tested on the same data set. The rate of data is extremely high, for example, around 2 GB per second. Clearly, this requires a large available storage space. Therefore, typically, the test drives are performed in the vicinity of main computing system 202 at the headquarters.

Main computing system 202 includes a first storage device 112 configured to store incoming raw data from a test vehicle 2 that is input via a data input device 110. First storage device 112 and the associated first computing device 114 may be part of a compute node or server group 113 in compute cluster 131. Likewise, a second storage device 116 and an associated second computing device 118 may be part of a second compute node or server group 117 in compute cluster 131. Each server node may be characterized by the ratio of computing power (number and quality of CPU cores and main memory) to its storage capacity (number and size of hard disks). For example, first server node 113 may have a relatively high computing power for data that is most likely accessed within the near future ("hot" data). Accordingly, the measurement data ingested by compute cluster 131 in the above-described manner is first stored on first storage device 112.

Second server node 117 may have an intermediate amount of computing power for data that might be accessed not in the immediate future, but perhaps in

the foreseeable future, or perhaps less frequently than the “hot” data (referred to herein as “warm” data). It will be appreciated that additional server nodes for data that is even less likely to be accessed, having even less computing power and considerably higher storage capacity, may also be provided (for “cold” data). In addition, an object store 140 that has practically no computing power is provided for data that is outdated, but has to be kept for various reasons (“frozen” data). It should be noted that, in some embodiments, at least some of the nodes having data with different temperatures may also be provided at geographically different locations, instead of being co-located with each other, for example, at the headquarters.

Data distribution device 120 is configured to classify the measurement data stored on the respective storage devices into data having different priorities, for example, based on one or more predetermined criteria, and to transfer data having a low priority to a server node that has lower computing power. For example, data distribution device 120 may be configured to classify some measurement data stored on first storage device 112 as having a lower priority and transfer the same to second storage device 116. In a similar manner, data that is stored on, for example, second storage device 116 may be transferred to first storage device 112, if necessary. Data classification can be based on, for example, access times, creation dates, or other meta data or content-related criteria. The DaSense software or any other appropriate software running on main computing system 202 may evaluate the stored data on a regular basis and automatically move data to the appropriate server group to achieve minimum average query times. For example, it may be desired that the storage device 112 is has a predetermined fill rate of, for example, 90 % at all times, such that as much data as possible can be processed by computing device 114 having the highest computational power. The skilled person will appreciate that there are many possibilities for the criteria that are used to assign a higher priority to certain data. For example, the criterion may be the age of the data, i.e., the most recent data

may have the highest priority, and older data may be transferred from storage device 112 to storage device 116. Of course, meta data indicating the type and location of all data may always be stored on storage device 112 to minimize query times. In this respect, it should be noted that, when data is moved to object store 140, only its meta data is directly accessible. Therefore, the data that is moved to the object store is specifically analyzed and indexed by data distribution device 120 to allow efficient recovery, if necessary, such that only a minimum amount of data has to be transferred from the object store to, for example, storage device 116. To this end, data management device 120 includes a meta data generation unit 136 configured to analyze and index the data to be stored in object store 140 and generate appropriate meta data. For example, the meta data of the test drives may be provided with tags describing, e.g., the time of day of the drive, weather conditions, the type of road, tunnels, special events such as accidents, traffic jams or the like.

Similarly to what has been described for the first embodiment, upon receipt of a query via query input device 124, analysis device 126 communicates with data management device 122 to determine the location of the query-relevant data, i.e., on which server group or storage device the data is located, and simultaneously performs the analysis on the relevant storage devices by the associated computing devices, i.e., the cluster compute cores of the associated compute node. In this respect, the DaSense software or any similar software may offer the possibility to execute code in parallel on an arbitrary number of cluster compute cores. In other words, several instances may be executed on a single cluster node. This may be realized in the following manner. For example, a container such as a Docker image (www.docker.com) may be created that contains the required operating system and all necessary packages. This container is sent to each node as a single initial step. Via Hadoop streaming or Spark streaming or any other appropriate job control, a Docker container is started, and newly developed code is injected and compiled in the container. Inside the

container, the new executable is started, possibly with runtime parameters specifying, for example, input data, configuration parameters and the like. Each time the execution is finished, the Docker container reports back to the Hadoop streaming job whether it was successful or not, and also reports any results. Optionally, the executable may be started again with different parameters. The whole process can be accessed via, for example, the DaSense software. In this manner, new algorithms can be tested on many more data sets than in the case where algorithm test runs are done on workstations, with every engineer having his own personal set of test data. According to the present embodiment, parameter tests and optimizations can be performed in parallel by evaluating one parameter set per instance with a minimum of programming effort.

Although preferred embodiments of the present invention have been described herein, it may be readily apparent that many modifications may be made to the disclosed data analysis system and method without departing from the scope of the appended claims. Accordingly, the specific embodiments described herein are not intended to limit the scope of the claims to the described embodiments.

It is explicitly stated that all features disclosed in the description and/or the claims are intended to be disclosed separately and independently from each other for the purpose of original disclosure as well as for the purpose of restricting the claimed invention independent of the composition of the features in the embodiments and/or the claims. It is explicitly stated that all value ranges or indications of groups of entities disclose every possible intermediate value or intermediate entity for the purpose of original disclosure as well as for the purpose of restricting the claimed invention, in particular as limits of value ranges.

Claims

1. A distributed data analysis system (100; 200) for analyzing collected measurement data, comprising:
 - a data input device (10; 110) configured to receive measurement data;
 - a first storage device (12; 112) associated with the data input device (10; 110) and configured to store the measurement data input via the data input device (10; 110);
 - a first computing device (14; 114) associated with the first storage device (12; 112);
 - a second storage device (16; 116) configured to store measurement data previously stored on the first storage device (12; 112);
 - a second computing device (18; 118) associated with the second storage device (16; 116);
 - a data distribution system (19; 119) configured to distribute the measurement data between the first storage device (12; 112) and the second storage device (16; 116) based on at least one predetermined criterion;
 - a data management device (22; 122) configured to store a location of the measurement data and to update the stored location based on the distribution by the data distribution device (20; 120);
 - a query input device (24; 124) configured to receive a query for an analysis to be performed on the collected measurement data;
 - an analysis device (26; 126) configured to perform the analysis on the measurement data stored on the first storage device (12; 112) and the second storage device (16; 116) by the first computing device (14; 114) and the second computing device (18; 118), respectively, based on the location of the collected measurement data stored by the data management device (22; 122); and

a reporting device (28; 128) configured to report a result of the analysis of the collected measurement data.

2. The system of claim 1, comprising:

a plurality of first storage devices (12) located at geographically different locations, each first storage device (12) being associated with a corresponding data input device (10) and a corresponding first computing device (14), wherein the second storage device (16) is a central storage device configured to store data previously stored on the plurality of first storage devices (12).

3. The system of claim 2, wherein

each first computing device (14) is configured to acquire a type of the measurement data input via the corresponding data input device (10), for example, by automatically detecting a data format of the same, and to perform a pre-processing of the measurement data based on the acquired type of measurement data prior to storing the measurement data on the associated first storage device (12).

4. The system of claim 3, wherein the pre-processing includes converting the measurement data into a format that is suitable for analysis in a cluster computing framework.

5. The system of any one of claims 2 to 4, wherein

each first computing device (14) is configured to generate meta data from the measurement data input via the corresponding data input device (10) and stored on the first storage device (20), said meta data including the location of the measurement data, and to forward the meta data to the data management device (22).

-25-

6. The system of any one of claims 2 to 5, wherein each first computing device (14) includes a computing cluster (30), for example, a Hadoop cluster, configured to perform the analysis of the measurement data stored on the first storage device (12).

7. The system of any one of claims 2 to 6, wherein the at least one criterion includes at least one of:
a remaining capacity of one of the plurality of first storage devices (12);
an age of the measurement data;
lapse of a predetermined time interval; and
a specific type of the measurement data stored on one of the plurality of first storage devices (12), and
the data distribution system (19) includes a data distribution device (20) configured to initiate a backup of at least part of the measurement data stored on one of the plurality of first storage devices (12) to a portable storage device (32) provided at the location of the corresponding first storage device (12) when the at least one criterion is met.

8. The system of claim 7, further comprising a backup data input device (34) provided at the location of the second storage device (16), connectable to the portable storage device (32) and configured to read the measurement data on the portable storage device (32) and to transfer the same to the second storage device (16).

9. The system of any one of claims 2 to 8, further comprising a data verification device (36) configured to verify that data that has been transferred to the second storage device (16) is identical to the measurement data previously stored on the first storage device (12), and to initiate deletion of the measurement

data previously stored on the first storage device (12) after successful verification of the data, wherein

the data management device (22) is configured to update the location of the verified measurement data after successful verification.

10. The system of any one of claims 2 to 9, wherein each of the plurality of first storage devices (12) is contained in a housing (40) together with the associated first computing device (14), the housing (40) including the data input device (10) and forming a standalone unit.

11. The system of claim 1, wherein
the second storage device (116) is co-located with the first storage device (112),
the first computing device (114) has more computing power than the second computing device (118), and
the data distribution device (120) is configured to classify the measurement data into data having different priorities, and to transfer data having a lower priority to the second storage device (116).

12. The system of claim 11, wherein the first computing device (114) and the second computing device (118) form part of a cluster computing system (119), and the data management device (120) is configured to classify the measurement data based on at least one of: access times; creation dates; types of data; and other meta data associated with the collected measurement data.

13. The system of claim 11 or 12, further comprising an object store (140) having substantially no computing power and being configured for long-term storage of measurement data having a lowest priority, wherein the data management device (120) includes a meta data generation device (136)

configured to analyze the data to be stored in the object store (40), and to generate appropriate meta data for efficient recovery of the stored measurement data by the data management device (120) based on the analysis.

14. The system of any preceding claim, wherein the analysis device (26; 126) is configured to generate analysis code to be executed on the first computing device (14; 114) and the second computing device (18; 118), respectively, based on the query, and to send the analysis code to the first computing device (14; 114) and the second computing device (18; 118) for execution on the same.

15. A method for distributed data analysis of collected measurement data stored on a first storage device (12; 112) and a second storage device (14; 114), comprising:

receiving measurement data;

storing a location of the received measurement data;

distributing the measurement data between the first storage device (12; 112) and the second storage device (16; 116) based on at least one predetermined criterion;

updating the location of the measurement data based on the distribution;

receiving a query for an analysis to be performed on the collected measurement data;

performing the analysis on the measurement data stored on the first storage device (12) and the second storage device (14; 114) by a first computing device (14; 114) associated with the first storage device (12; 112) and a second computing device (18; 118) associated with the second storage device (16; 116), respectively, in accordance with the stored location of the measurement data; and reporting a result of the analysis.

16. The method of claim 15, further comprising
acquiring a type of the received measurement data, and
performing pre-processing of the measurement data based on the acquired
type of measurement data prior to storing the measurement data.

17. The method of claim 16, wherein pre-processing includes
converting the measurement data into a different format, for example, a format
that is suitable for analysis in a cluster computing framework.

18. The method of any one of claims 15 to 17, further comprising
generating meta data from the received measurement data, said meta data
including the location of the measurement data.

19. The method of any one of claims 15 to 18, further comprising
initiating a backup of at least part of the measurement data stored on the
first storage device (12) to a portable storage device (32) provided at the location
of the corresponding first storage device (12) when the at least one criterion is
met, the at least one criterion including at least one of:

- a remaining capacity of the first storage device (12);
- an age of the measurement data;
- lapse of a predetermined time interval; and
- a specific type of the measurement data stored on the first storage device
(12).

20. The method of claim 19, further comprising
physically transporting the portable storage device (32) to the location of
the second storage device (16), for example, by mail,
transferring the backup data from the portable storage device (32) to the
second storage device (16),

verifying that the transferred data is identical to the data previously stored on the first storage device (12), and

deleting the measurement data previously stored on the first storage device after successful verification.

21. The method of claim 15, further comprising
classifying the measurement data into data having different priorities based on at least one of access times, creation dates, types of data, and other meta data associated with the collected measurement data, and
transferring data having a lower priority from the first storage device (112) to the second storage device (116).

22. The method of any one of claims 15 to 21, further comprising
generating analysis code based on the query, and
executing the analysis code on the first computing device (14; 114) and the second computing device (18; 118).

23. A computer program comprising computer-executable instructions that, when executed on a computer system, cause the computer system to execute the steps of the method of any one of claims 15 to 22.

24. A computer program comprising computer-executable instructions that, when executed on a computer system, cause the computer system to perform the following steps:

acquiring a location of measurement data stored on a plurality of storage devices (12, 16) provided at geographically different locations;

receiving a query for an analysis to be performed on the measurement data;

-30-

evaluating the query to determine on which of the plurality of storage devices (12, 16) data relevant to the query is stored;
generating analysis code based on the query;
forwarding the generated analysis code to computing devices (14, 18) associated with the relevant storage devices (12, 16), respectively;
receiving a partial analysis result from each computing device (14, 18);
combining the partial analysis results; and
reporting the combined analysis result.

25. The computer program of claim 24, further comprising instructions for initiating a transfer of measurement data from one storage device to another when at least one predetermined criterion is met, the at least one criterion including at least one of:

a remaining capacity of the first storage device (12);
an age of the measurement data;
lapse of a predetermined time interval; and
a specific type of measurement data stored on the first storage device (12).

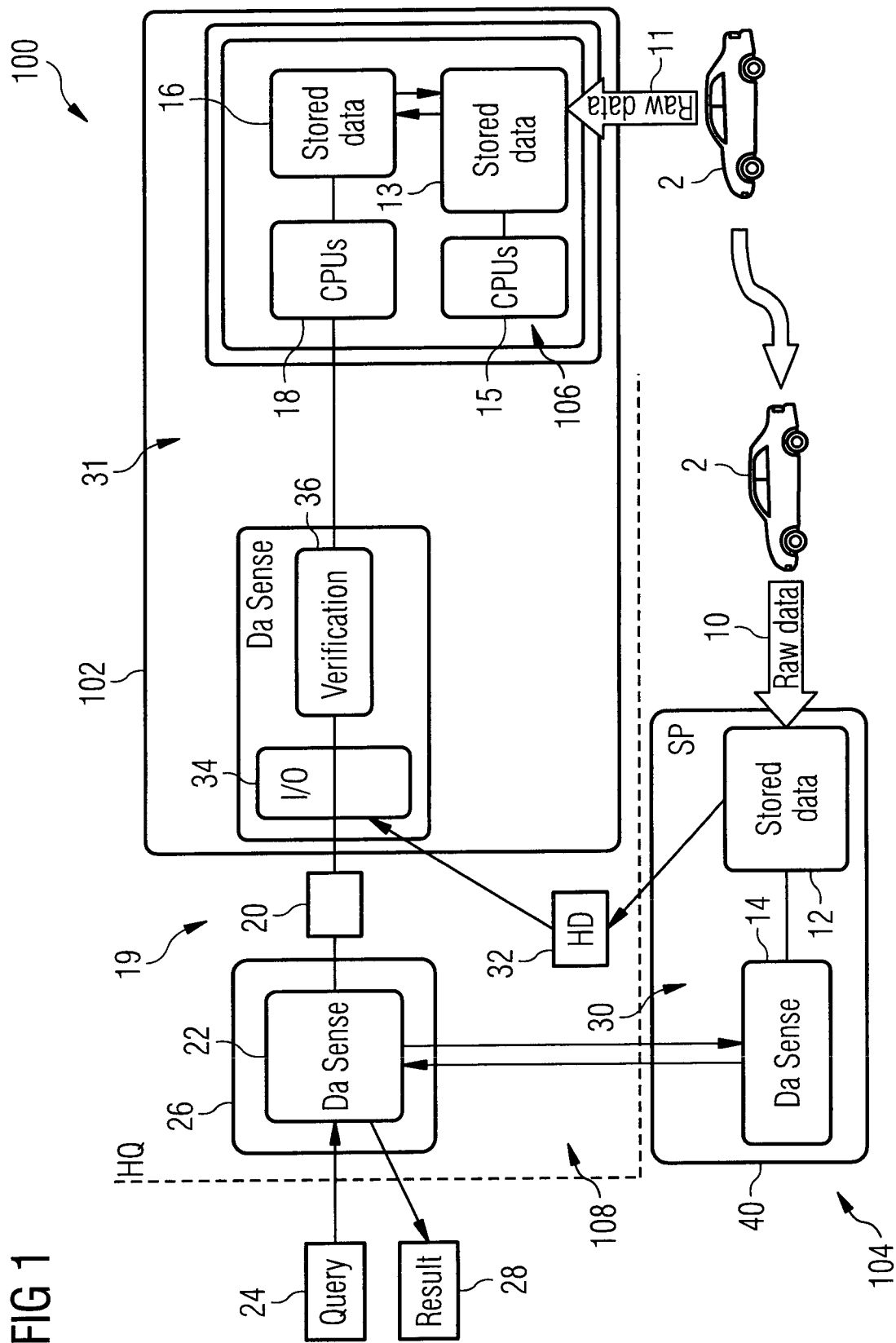
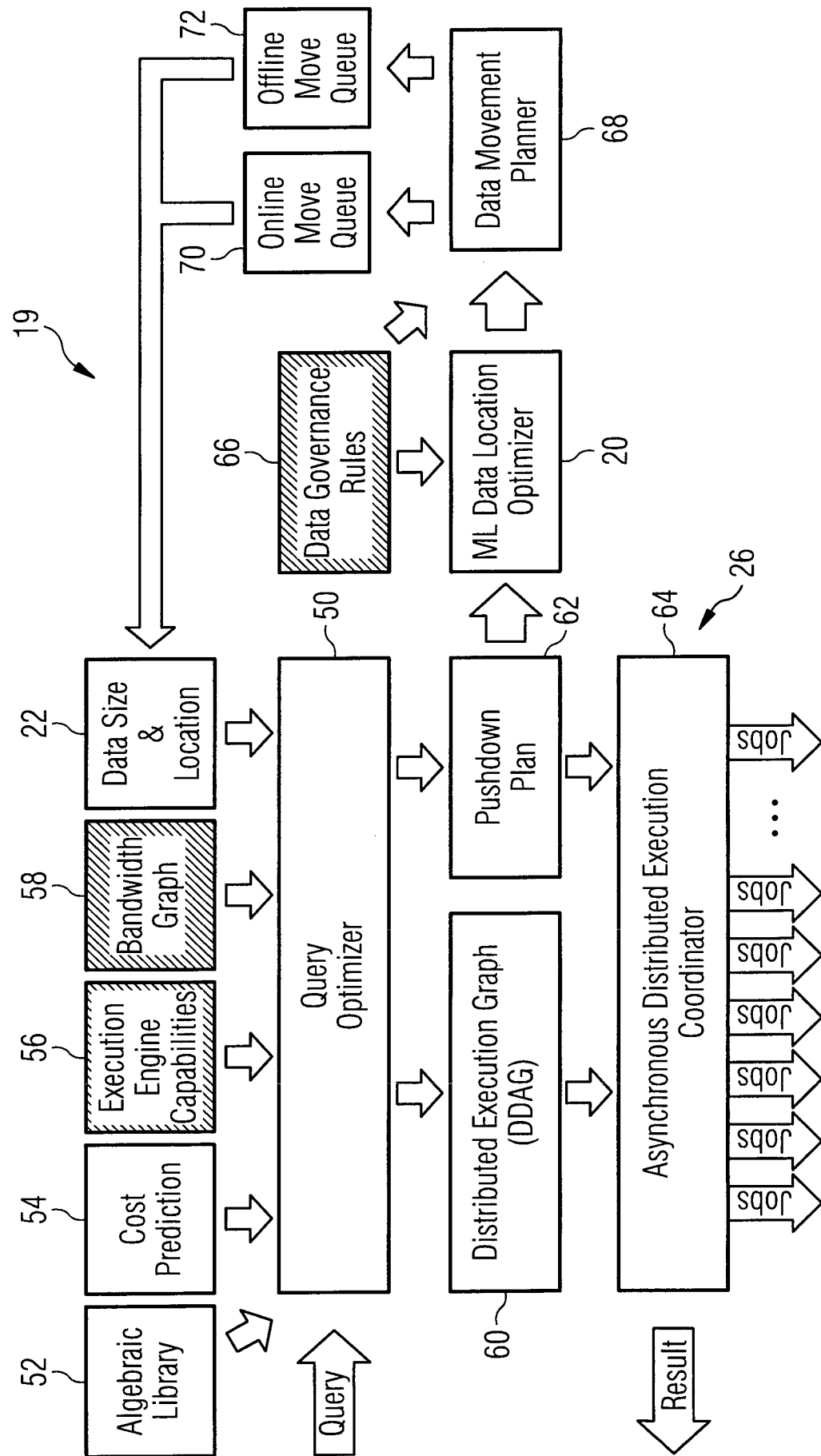
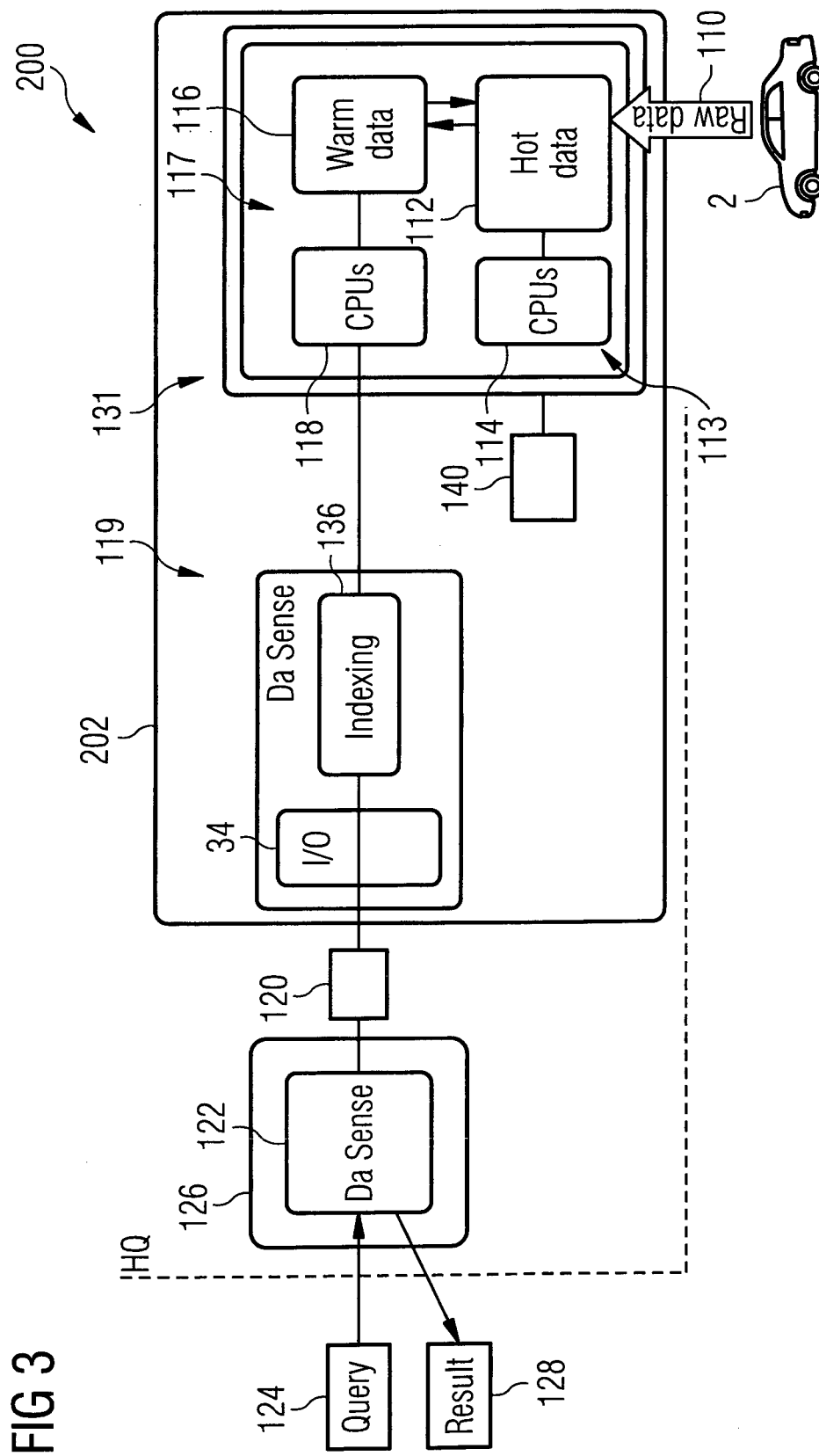


FIG 2





INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2016/000713

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06F9/50

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014/040575 A1 (HORN GUSTAV [US]) 6 February 2014 (2014-02-06) figures 1-4,7 paragraphs [0008], [0009], [0023], [0024], [0027] - [0031], [0035] - [0039], [0046], [0047] -----	1-25
X	US 2014/195558 A1 (MURTHY RAGHOTHAM [US] ET AL) 10 July 2014 (2014-07-10) figures 1-5 paragraphs [0004] - [0010], [0022], [0029], [0035], [0059] -----	1-25



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 June 2016

Date of mailing of the international search report

22/06/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Michalski, Stéphane

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/000713

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2014040575 A1	06-02-2014	US 2014040575 A1	06-02-2014
		WO 2014022674 A1	06-02-2014

US 2014195558 A1	10-07-2014	AU 2013371448 A1	23-07-2015
		CA 2896855 A1	10-07-2014
		CN 104903894 A	09-09-2015
		EP 2752779 A2	09-07-2014
		JP 2016509294 A	24-03-2016
		KR 20150104585 A	15-09-2015
		US 2014195558 A1	10-07-2014
		US 2015261831 A1	17-09-2015
		WO 2014107359 A1	10-07-2014
