



(51) International Patent Classification:
G05B 19/418 (2006.01)

(21) International Application Number:
PCT/RU2018/000420

(22) International Filing Date:
25 June 2018 (25.06.2018)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **SIEMENS AKTIENGESELLSCHAFT**
[DE/DE]; Werner-von-Siemens-Strasse 1, 80333 Munchen
(DE).

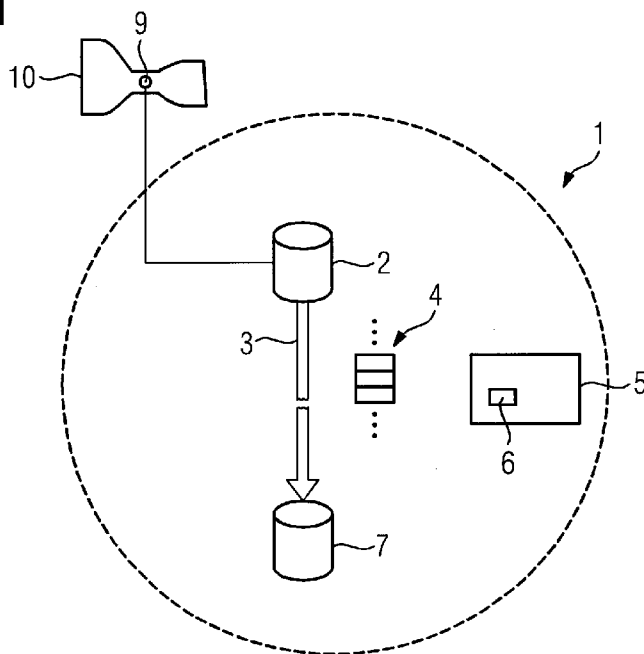
(72) Inventors: **GOLOSHCHAPOV, Egor Sergeevich**;
Engelsa street, 124/3, apt.7, St.Petersburg, 194356
(RU). **OZHIGIN, Artem Vladimirovich**; Pogranichnika

Garkavogo, 26/4-10, St.Petersburg, 198264 (RU). **VINO-GRADOV, Sergey Valeryevich**; Rimskogo-Korsakova str., 109-22, St.Petersburg, 190121 (RU). **TCYMBAL, Alexey Yurevich**; Nastavnikov, 3/1, apt. 15, St.Petersburg, 195298 (RU). **ZOBNIN, Sergey Sergeevich**; Pulkovskoe highway, 26/7, apt. 308, St.Petersburg, 196158 (RU).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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.

(54) Title: SYSTEM AND METHOD FOR STREAMING MEASUREMENT DATA

FIG 1



(57) Abstract: A system for streaming measurement data comprises: a number of provider devices configured to provide a number of data streams each including multiple structured data records indicative of measurement data obtained from a number of data sources of an industrial facility; a registry device configured to store, for each data stream, a semantics specification indicative of a semantics of the structured data records of the data stream; and a number of consumer devices configured to, for at least one data stream: receive the data stream provided by the corresponding provider device; obtain the measurement data from the structured data records based on the semantics specification of the received data stream stored in the registry device; and consume the obtained measurement data. Versatility is increased when multiple provider devices operate according to different semantics. A corresponding method and computer program product are also proposed.



(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, 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))*

SYSTEM AND METHOD FOR STREAMING MEASUREMENT DATA

The present invention relates to the field of industrial sensor data processing and more particularly to a system and method for streaming measurement data.

5 In a modern industrial facility, such as a power plant, gas turbine, processing plant, and the like, a large number of data sources such as sensors are installed. A respective provider device may provide measurement acquired by a respective data source to a streaming platform implemented in
10 a cloud or data center. The streaming platform may aggregate, process and/or store data streams of the measurement data. A number of consumer devices may consume the data streams of the streaming platform to perform analysis, monitoring, control and the like of the industrial facility. A respective data
15 stream may be a real-time, high-frequency data stream.

A problem exists in that in an industrial scenario, different provider devices may use different data representations for the data stream provided to the streaming platform. Likewise, different consumer devices may expect different data
20 representations in the data streams consumed from the streaming platform. For example, when using a commercially available real-time data streaming platform such as Apache Kafka or Confluent I/O, additional effort is required to provide interfaces between respective data sources, provider
25 devices, consumer devices, and the streaming platform. Tedious and error prone development is required each time a new provider or consumer device is added to the system.

It is one object of the present invention to provide an improved system and method for streaming industrial high-
30 frequency measurement data with increased versatility.

Accordingly, a system for streaming measurement data comprises: a number of provider devices each configured to

provide one or more of a number of data streams each including a plurality of structured data records indicative of measurement data obtained from one of a number of data sources of an industrial facility; a registry device configured to store, for each of the number of data streams, a semantics specification indicative of a semantics of the plurality of structured data records of the respective data stream; and a number of consumer devices configured to, for at least one of the number of data streams: receive the respective data stream provided by the corresponding provider device; obtain the measurement data from the plurality of structured data records included in the received data stream based on the semantics specification of the received data stream stored in the registry device; and consume the obtained measurement data of the received data stream.

The proposed system may advantageously enable a respective consumer device to consume a variety of data streams structured according to a variety of semantics. The system may employ a variety of different provider devices and consumer devices while obviating a necessity to install an interface component on a respective provider and/or consumer device.

"Streaming", as used herein, may refer to one or more of transmitting, receiving and processing, in particular processing in real-time, of a number of data streams.

The term "a number of", as used herein, shall mean a number of one or more.

The proposed system may comprise a streaming platform. The streaming platform may comprise a number of computing resources. The number of computing resources may include at least the number of provider devices, the number of consumer devices the registry device, and zero, one or more further computing resources. The number of computing resources of the streaming platform may be interconnected using a network, such

as a local area network (LAN), a wide area network (WAN), and the like.

The streaming platform may be configured to provide one or more of the following functionalities: enabling a provider
5 device to provide a data stream to the streaming platform; enabling a consumer device to consume a data stream from the streaming platform; temporarily or permanently storing a data stream; recovering a stored data stream; executing a streaming application on a data stream. A streaming application executed
10 on a data stream may be configured to alter or transform the respective data stream and/or to perform an action such as monitoring, analysis, control and the like, of the industrial facility based on the data stream.

As used herein, "provide a data stream" may mean directly
15 transmitting the data stream to one of the number of consumer devices consuming the data stream and/or may mean transmitting the data stream to another computing resource of the streaming platform for further processing.

Conversely, "receive a data stream provided by a provider
20 device" may mean directly receiving the data stream from the provider device and/or may mean receiving the data stream and/or a processed, altered or modified version thereof, from another computing resource of the streaming platform.

In other words, either no or any number of computing resources
25 of the streaming platform may be interposed between a provider device and a consumer device.

In particular, the received data stream may have been altered by processing that has been carried out one or more of the number of computing resources of the streaming platform on the
30 data stream provided by the provider device.

A respective data source may be configured to repeatedly acquire a measurement value from the industrial facility. A

respective data source may be a sensor, such as a temperature sensor, pressure sensor, and the like.

A respective provider device and a corresponding data source may be separate units connected to each other or may be embodied as a respective integral unit.

A respective provider device may be connected to, or formed integrally with, any number of data sources and may provide any number of data streams to the streaming platform.

A data stream may be a signal or a sequence of data packets in transmission. A data stream may comprise a plurality of structured data records. Each structured data record may comprise a representation of a respective measurement value as one piece of measurement data and may further comprise zero or any number of further representations of further other pieces of measurement data such as a timestamp, a manufacturer code, and the like. Each structured data record may be formatted in a format such as JSON, XML, or the like.

"Semantics", herein, may refer to a data representation or information that can be used to retrieve a piece of measurement data, such as a measurement value, from a respective structured data record. For example, "semantics" may refer to a respective field name or identifier of a field containing a representation of a respective piece of the measurement data. A "semantics specification" may refer to a representation of a semantics in a machine-readable form.

The semantics specifications of the number of data streams may be, for example, pre-stored in the registry device. When a further provider device is added to the system, an engineer may register a corresponding further semantics specification in the registry device without having to alter the further provider device, install an interface component on the further provider device or the like.

The phrase "obtain the measurement data based on the semantics specification" may refer to accessing a respective specific field of a respective structured data record, wherein a respective specific field is a field that contains the representation of a respective piece (portion) of the measurement data, and reading the respective piece of the measurement data, such as the measurement value, from the specific field, wherein the respective specific field is identified among a plurality of fields comprised by the structured data record based on the semantics specification.

According to one variant, a respective consumer device may receive the semantics specification from the registry device so as to be able to determine the respective specific field based on the semantics specification.

According to another variant, another computing resource may receive the semantics specification from the registry device, may determine the respective specific field based on the semantics specification, and which may add a respective tag or the like to the specific field in the structured data record of the data stream. A respective consumer device may access the respective specific field by locating the respective tag in the structured data record of the data stream.

Both variants shall be encompassed by the phrase "obtaining the measurement data based on the semantics specification".

"Consuming the obtained measurement data" may refer to any processing that is carried out using the obtained measurement data as input data. In particular, "consuming the obtained measurement data" may refer to performing processing that has a technical effect, such as monitoring, analysis or control of the industrial facility, based on the obtained measurement data. Said processing may be performed by the consumer device, by another computing resource of the streaming platform

connected to the consumer device, or by an external unit connected to the consumer device.

It is noted that "obtaining the measurement data" and "consuming the obtained measurement data" may also include a case where "at least portions of the measurement data" are obtained and/or consumed and not all of the measurement data included in the data stream is obtained and/or consumed.

In the proposed system, the semantics specifications stored in the registry device may advantageously be used to make the measurement data indicated in the structured data records available to a respective consumer device. Herein, a provided device may use any semantics, and, more particular, is not required to use a pre-agreed semantics. Thereby, versatility of the data streaming system may be increased in cases where an industrial data streaming application is built from a heterogenous plurality of data sources and/or provider devices.

According to one embodiment, the registry device is configured to provide an interface for registering one or more of the semantics specifications.

The interface may be an interface for automated interaction, such as an Application Programming Interface (API) or a web interface, and/or may be an interface for human interaction such as a display and an input device such as a keyboard, mouse, touch panel or the like.

Thereby, a versatility when registering new provider devices when expanding the system may be increased.

According to a further embodiment, the registry device is configured to determine one or more of the semantics specifications by parsing semantics information about the corresponding data stream, said semantics information being received from the corresponding provider device.

The "semantics information" may be any information provided by the corresponding provider device of the data stream that enables the registry device to determine a semantics specification for the data stream. For example, the semantics
5 information may be a textual specification of the structure of the structured data records comprising, for example, names of fields and their meanings and the like.

That is, the registry device may determine a respective semantics specification automatically through interaction with
10 the provider device. An engineer building or expanding the system may be favorably relieved from the burden of registering respective semantics specifications for each provider device with the registry device.

According to a further embodiment, the semantics information
15 received by the registry device comprises at least one of the plurality of structured data records of the corresponding data stream.

Herein, in order to automatically determine the semantics information, the registry device may subscribe to a data
20 stream provided by a respective provider device. That is, the semantics specification may be determined by parsing a sample of the actual data stream.

Thereby, favorably, the registry device may automatically determine a respective semantics specification without special
25 support by a corresponding provider device, solely based on the data stream provided by the provider device. The registry device may use a heuristic approach, artificial intelligence or the like to automatically determine the respective semantics specification.

30 According to a further embodiment, a respective of the plurality of structured data records comprises a number of fields, each field including a value and a field identifier

associated with the value, and a respective of the number of semantics specifications comprises a mapping between each of a number of predetermined tags and a corresponding one of the field identifiers.

5 It is noted that the number of fields may be greater than one. That is, apart from a field including the measurement value as a value, a structured data record may include further fields including further values such as a timestamp, a version number, a sensor identifier, a manufacturer identifier and the
10 like.

The number of predetermined tags may be understood as a syntax predefined in the proposed system. A tag, and/or a respective instance thereof, may refer to a string or textual label such as "<sensor_value>", "<timestamp_value>", "<sensor_name>" and
15 the like. A respective tag may be associated with a corresponding piece of measurement data, such as a measurement value, a timestamp value, a sensor name, and the like.

The mapping may be an array, a hash table or any other machine-readable data structure allowing to determine a
20 corresponding field identifier for a given predetermined tag.

A device such as the consumer device, a tagging device or the like, which is interested in determining, among the number of fields included in a structured data record of a given data stream, a specific field including a specific piece of
25 measurement data, may obtain the semantics specification of the data stream from the registry device, may query the mapping included in the semantics specification using a known predetermined tag associated with the specific piece of measurement data, such as "<sensor_value>", may obtain, as a
30 response to the query, the field identifier, among the plurality of field identifiers, which is associated to the specific field including the specific piece of measurement data.

Thereby, advantageously, the consumer device may not be required to know a specific semantics of a data stream provided by a provider device, but it may be sufficient to know the syntax or semantics of the predetermined tags. This
5 may advantageously allow to increase reusability of code implemented on the consumer device for data streams provided by a variety of heterogenous provider devices with different semantics.

According to an embodiment, said parsing of the received
10 semantics information includes matching each of a number of patterns defined for each of the number of predetermined tags against each of a number of field identifiers included in the received semantic information.

A respective pattern may be a wildcard pattern, a regular
15 expression, or the like.

Thereby, advantageously, common portions of field identifiers that are commonly found in data streams provided by data sources and/or provider devices manufactured by different manufacturers, such as the string "*value*", may be pre-stored
20 in the registry device and may be used to automatically determine a respective semantics specification.

According to a further embodiment, at least one of the number of consumer devices is further configured to receive the semantics specification of the respective data stream from the
25 registry device.

That is, for example, rather than hardcoding specific field names used in the consumer device, the consumer device may be configured to communicate with the registry device to obtain, for a given data stream, the corresponding field name based on
30 a predetermined tag.

Reusability of the consumer device for various data streams provided by various provider devices may be favorably enhanced.

According to another embodiment, the system for streaming
5 measurement data further comprises a tagging device configured to, for at least one of the number of data streams: receive, from the registry device, the semantics specification of the respective data stream; receive the respective data stream provided by a corresponding one of the provider devices;
10 modify the received data stream by inserting, based on the received semantics specification, instances of one or more of the number of predetermined tags into the plurality of structured data records of the modified data stream; and forward the modified data stream to each of the number of
15 consumer devices consuming the corresponding data stream.

The tagging device may be one of the number of computing resources of the streaming platform. The tagging device and the registry device may be embodied as an integrated or same unit or may be embodied as separate units connected to each
20 other.

That is, a respective data stream may travel, directly or via a number of further computing resources, from the provider device to the tagging device. The tagging device may insert instances (copies) of one or more of the predetermined tags
25 into each of the plurality of structured data records in accordance with the received semantics information of the respective data stream. The modified data stream including the tagged structured data records may then travel from the tagging device, directly or via a number of further computing
30 resources, to the consumer device.

The instances of the predetermined tags in the modified data stream may aid the consumer device in locating pieces of interested of the measurement data, such as a measurement

value, in the structured data records of the modified data stream without requiring knowledge of the detailed semantics of the (modified) data stream.

According to a further embodiment, the tagging device is
5 configured to, for each field of each of the plurality of structured data records, if it is determined, based on the semantics specification, that the field identifier of the respective field corresponds to one of the predetermined tags, associate an instance of the corresponding one of the
10 predetermined tags with the field identifier.

Associating an instance of a predetermined tag with a field identifier may refer to placing the instance of the predetermined tag in positional association with the field identifier, such as next to, before, after, or near to the
15 field identifier.

The tagging device may thus flag each field identifier associated with a field including a piece of measurement data which is known in the pre-defined syntax of the proposed system by placing an instance of a corresponding predetermined
20 tag.

Thereby, the recognition of individual pieces of measurement data in the structured data records of the modified data stream is facilitated.

According to a further embodiment, a respective one of the
25 number of consumer devices is configured to obtain the measurement data from the plurality of structured data records by accessing a number of fields of a respective structured data record, each of the accessed number of fields having an associated field identifier that is associated with one of the
30 predetermined tags.

That is, the consumer device may locate a predetermined tag corresponding to a piece of measurement data of interest in

the structured data record, may thereby identify, out of a plurality of fields comprised by the structured data record, a field next to, near to, before or after the instance of the predetermined tag, and may read the value of the identified
5 field to obtain the desired piece of measurement data. This procedure may be repeated for each piece of measurement data for which a corresponding predetermined tag is defined in the proposed system until the entire measurement data (or each piece of interest thereof) is obtained.

10 That is, the consumer may favorably neither require prior knowledge about a semantics of the data stream. Further, a necessity for the consumer device to communicate with the registry device to obtain a corresponding semantics specification is obviated. The consumer device may access
15 pieces of measurement data of interest easily by referring to the instances of the predetermined tags inserted into the data stream.

According to a further embodiment, consuming the obtained measurement data includes providing the obtained measurement
20 data to an external unit connected to the system for streaming measurement data.

That is, the consumer device may comprise an interface device for coupling a third-party device to the proposed system and providing the obtained measurement data to the third-party
25 device according to a predetermined format.

According to a further embodiment, consuming the obtained measurement data includes storing the obtained measurement data into a historical database comprised by the system for streaming measurement data.

30 It is noted that the obtained measurement data may only comprise pieces of measurement data for which a corresponding predetermined tag is defined in the proposed system. Other

pieces of measurement data which may also be included into a data stream by some of the provider devices but which do not correspond to any of the predetermined tags defined in the proposed system may not be included in the obtained measurement data. Therefore, by storing only the obtained measurement data obtained from a respective of the data streams, an amount of storage space of the historical database used for storing the measurement data may be reduced in comparison to storing the entire data stream provided by the provider device.

According to a further embodiment, consuming the obtained measurement data includes performing, on the consumer device and/or a number of further devices connected to the consumer device and comprised by the system for streaming measurement data, at least one of analysis, monitoring, and control of the industrial facility using the obtained measurement data.

Versatility of analysis, monitoring and control of the industrial facility may be favorably improved by allowing a variety of data sources using different semantics, to be added to the proposed system without requiring a special configuration, interface or the like to be implemented on the respective data sources and/or their corresponding provider devices.

According to a further embodiment, the measurement data is high-frequency measurement data obtained from the number of data sources with a frequency of 1 kHz or higher.

It is noted that tagging a high-frequency data stream with a frequency of 1 kHz, more preferably 10 kHz, or greater may require a considerable amount of computing power. Therefore, according to the present embodiment, the tagging device may be preferably implemented by a plurality of computing resources of a distributed high-performance streaming platform.

According to a further embodiment, number of data streams comprise a number of real-time data streams provided by the number of provider devices and received by the number of consumer devices in real time.

- 5 In particular, the tagging device may be configured to perform said tagging on a live data stream in real time.

Thereby, the above-mentioned advantages associated with tagged data streams may be made available to real-time analysis and/or control processing performed by a high-performance,
10 real-time, data-center or cloud-based streaming platform.

An engineer tasked with developing such real-time analysis and/or control processing may be relieved from the burden of dealing with various differing semantics used by various of the provider devices.

- 15 The respective entity, e.g. a respective provider device, the registry device, the tagging device, and a respective consumer device, may be implemented in hardware and/or in software. If said entity is implemented in hardware, it may be embodied as a device, e.g. as a computer or as a processor or as a part of
20 a system, e.g. a computer system. If said entity is implemented in software it may be embodied as a computer program product, as a function, as a routine, as a program code or as an executable object.

Any embodiment of the first aspect may be combined with any
25 embodiment of the first aspect to obtain another embodiment of the first aspect.

According to a second aspect, a method for streaming measurement data comprises: providing, by each of a number of provider devices, one or more of a number of data streams each
30 including a plurality of structured data records indicative of measurement data obtained from one of a number of data sources of an industrial facility; storing, in a registry device, for

each of the number of data streams, a semantics specification indicative of a semantics of the plurality of structured data records of the respective data stream; and performing, by a number of consumer devices and for at least one of the number
5 of data streams, the steps of: receiving the respective data stream provided by the corresponding provider device; obtaining the measurement data from the plurality of structured data records included in the received data stream based on the semantics specification of the received data
10 stream stored in the registry device; and consuming the obtained high-frequency measurement data of the received data stream.

The embodiments and features described with reference to the proposed system for streaming measurement data apply mutatis
15 mutandis to the proposed method for streaming measurement data.

According to a further aspect, a computer program product is proposed comprising a program code for executing the above-described method for streaming measurement data when run on at
20 least one computer.

A computer program product, such as a computer program means, may be embodied as a memory card, USB stick, CD-ROM, DVD or as a file which may be downloaded from a server in a network. For example, such a file may be provided by transferring the file
25 comprising the computer program product from a wireless communication network.

Further possible implementations or alternative solutions also encompass combinations - that are not explicitly mentioned herein - of features described above or below with regard to
30 the embodiments. The person skilled in the art may also add individual or isolated aspects and features to the most basic form of the invention.

Further embodiments, features and advantages will become apparent from the subsequent description and dependent claims, taken in conjunction with the accompanying drawings, in which:

Fig. 1 shows a system for streaming measurement data according to a first exemplary embodiment;

Fig. 2 shows a method for streaming measurement data according to the first exemplary embodiment;

Fig. 3 shows a system for streaming measurement data according to a second exemplary embodiment; and

Fig. 4 shows a system for streaming measurement data according to a third exemplary embodiment.

In the Figures, like reference numerals designate like or functionally equivalent elements, unless otherwise indicated.

Fig. 1 shows a system 1 and Fig. 2 shows a corresponding method for streaming measurement data according to the first exemplary embodiment.

The system 1 comprises a data provider device (provider device) 2, a registry device 5 and a data consumer device (consumer device) 7. The data provider device 2 is connected to a sensor 9 (example of a data source) installed in a gas turbine 10 (example of an industrial facility).

In step S1 of the method of the first exemplary embodiment, the data provider device 2 repeatedly acquires measurement data from the sensor 9 installed in the gas turbine 10 and creates a data stream 3 including a plurality of structured data records 4.

Each structured data record 4 is indicative of measurement data obtained during a respective acquisition and may comprise a plurality of pieces of measurement data such as a measurement value, a sensor name, a timestamp, and the like. Each structured data record 4 of the data stream 3 is

structured according to a semantics specific to the data provider device 2 and/or the sensor 9 (also referred to as "semantics of the data stream 3").

5 In step S2, a semantics specification 6 is stored in the registry device 5. The semantics specification 6 is indicative of the semantics of the data stream 3.

10 In step S3, the data consumer device 7 receives the data stream 3 that is provided by the data provider device 2. It is noted that the data consumer device 7 is not aware of the semantics of the data stream 3. That is, no information on the semantics of the data stream 3 is configured in the data consumer device 7.

15 Nevertheless, in step S4, the data consumer device 7 obtains at least portions of the measurement data included in the plurality of structured data records 4 included in the received data stream 3.

20 More specifically, said obtaining of the at least portions of the measurement data from the data stream 3 by the data consumer device 7, which is unaware of the semantics of the data stream 3, is enabled by the semantics specification 6 stored in the registry device 5.

In step S5, the obtained measurement data (at least the obtained portions thereof) included in the plurality of structured data records 4 of the data stream 3 is consumed.

25 It is noted that steps S1 through S5 are not required to be performed, and are typically not performed, in this particular order. Specifically, steps S1 and steps S3 to S5 may be executed in parallel to each other and may each be repeated in continuous loops, whereas step S2 is executed at least once
30 before a first execution of step S4.

According to the first exemplary embodiment, thanks to the semantics specification 6 stored in the registry device 5, the

data consumer device 7 is able to successfully consume at least portions of the measurement data included in the data stream 3 also in a case where there is no agreement between the data provider device 2 and the data consumer device 7 on a common semantics for the data stream 3.

The benefits of this technique may become more salient when considering that in a typical real-world scenario of using a streaming system 1 to process measurement data acquired in an industrial facility 10, there may typically be dozens, hundreds or even thousands of data sources 9 with corresponding data provider devices 2, which typically do not supply data according to a common semantics, but rather use varying semantics native to a respective of the data provider devices 2 and/or data sources 9.

According to one further development of the first exemplary embodiment, the data consumer device 7 consumes the obtained measurement data by performing processing for analysis, monitoring and/or control of the gas turbine 10 using the obtained measurement data as input data of said processing.

According to another further development of the first exemplary embodiment, the data consumer device 7 provides the obtained measurement data to an external unit (not shown in Fig. 1; 11 in Fig. 4) connected thereto.

Fig. 3 shows a system 1 for streaming measurement data according to a second exemplary embodiment.

The system of Fig. 3 may be a cloud-based streaming platform 1 and comprises at least four data provider devices 21, 22, 23, 24, two data consumer devices 71, 72 and a registry device 5. The streaming platform 1 may comprise an arbitrary number of interconnected further computing resources in addition to the devices (computing resources) 21-24, 71, 72 and 56.

Each of the data provider devices 21-24 provides a corresponding data stream 31-34 to the streaming platform 11. Each data stream 31-34 uses a specific semantics particular to the respective data stream 31-34 and/or data provider device
5 21-24.

Elements shown in Fig. 1 but not in Fig. 3 may nevertheless be present in the second exemplary embodiment, but are not illustrated for brevity. Specifically, although not shown in Fig. 3, also the data streams 31-34 include a respective
10 plurality of structured data records (4 in Fig. 1) indicative of measurement data obtained from a respective of a plurality of data sources (9 in Fig. 1) installed in an industrial facility (10 in Fig. 1).

The registry device 5 provides a programming interface for
15 registering semantics specifications 61-64 for each of the data streams 31-34 provided to the streaming platform 11. An engineer, when building and/or upgrading the system 1, may use an assistance device 12 to connect to the registry device 5 and store the semantics specifications 61-64 therein.

20 Each of the data consumer devices 71, 72 may receive any of the data streams 31, 32, 33, 34 present in the streaming platform 11.

When, for example, the second data consumer device 72 wants to receive the third data stream 33 provided by the third data
25 provider device 23, the second data consumer device 72 connects to the registry device 5 and asks for the semantics specification 63 of the third data stream 33.

Upon receiving the semantics specification 63 of the third data stream 33 from the registry device 5, the second data
30 consumer device 72 applies the received semantics specification 63 to obtain the measurement data included in the plurality of structured data records (4 in Fig. 1) included

in the received third data stream 33. The second data consumer device 72 then consumes the obtained measurement data.

Similar processing is carried out when any of the data consumer devices 71, 72 wants to receive any of the data streams 31, 32, 33, 34 provided by any of the data provider devices 21, 22, 23, 24, and a description thereof will not be repeated.

It is noted that a respective data stream 31-34 may pass through an arbitrary number of further computing resources (not shown) interposed between a respective data provider device 31-34 and a respective data consumer device 71-72.

Fig. 4 shows a system 1 for streaming measurement data according to a third exemplary embodiment.

The system of Fig. 4 may be a cloud-based streaming platform 1. The streaming platform 1 of Fig. 4 differs from the streaming platform of Fig. 3 in that the streaming platform 1 comprises a tagging device 8, a further computing resource 13, and in that the registry device 5 is not connected to an assistance device 12. These differences will now be described in more detail.

Specifically, the first data stream 31 provided by the first data provider device 21 may comprise a plurality of structured data records (4 in Fig. 1) using a first semantics as shown below:

```
sensorValue:  double
timestamp:    long
id:           string
```

Herein, "sensorValue", "timestamp" and "id" are field identifiers, and "double", "long" and "string" are data types of corresponding field values.

That is, for example, if the plurality of structured data records of the first data stream 31 are indicative of pressure values measured by a pressure sensor named "press", a sequence of two consecutive structured data records (4 in Fig. 1) of the first data stream 31 may be represented as follows:

```
sensorValue:  350.7
timestamp:    1528444828
id:           press1
sensorValue:  348.7
timestamp:    1528444829
id:           press1
```

The second data stream 32 may use the same semantics as the first data stream 31, but may provide a value of "press2" for the "id" field.

The third data stream 33 and the fourth data stream 34 provided by the third and fourth data provider devices 33, 34 may each comprise a plurality of structured data records using a second semantics as shown below:

```
manufacturer: string
sensorName:   string
temperature:  float
utime:        long
```

That is, one of the structured data records of the third data stream 33 may be represented as follows:

```
manufacturer: SIE
sensorName:   templ
temperature:  1321.8
utime:        1528444829
```

One of the structured data records of the fourth data stream 34 may be represented as follows:

```
manufacturer: SIE
sensorName:    temp2
5    temperature: 1139.3
    utime:      1528444829
```

It is noted that the above representations and formats are illustrative examples, and that the respective structured data records may also be formatted in a machine-readable format such as JSON, XML and the like.

Herein, in order to perform processing such as analysis, monitoring and control of the industrial facility (10 in Fig. 1), a respective data consumer device 71, 72 may be interested in certain pieces of measurement data of interest, such as a respective sensor value (pressure or temperature) and a respective sensor name. However, without further manual configuration and without support by the streaming platform 11, the consumers devices 71, 72 would be "blind" for the semantics of the respective data streams 31-34, i.e., the consumer devices 71, 72 would not know which of the fields of the respective structured data records contain the pieces of measurement data of interest.

Therefore, in the streaming platform 11, a predetermined syntax comprising the predetermined tags "<value>" and "<name>" is defined for (agreed upon by or known to) the plurality of data consumer devices 71, 72, the tagging device 8 and the registry device 5.

A respective semantics definition 61-64 stored in the registry device 5 comprises a mapping between the predetermined tags and the field identifiers of a corresponding data stream 31-34.

For example, given the exemplary structured data records of the various data streams 31-34 shown above, the semantics definitions 61 and 62 may each be represented as follows:

<value> : sensorValue

5 <name> : id

The semantics definitions 63 and 64 may each be represented as follows:

<value> : temperature

<name> : sensorName

10 In the present embodiment, for example, the tagging device 5 may be configured to automatically determine the semantics definitions 61 to 64.

More particularly, the tagging device 5 may acquire one of the plurality of structured data records of each of the data streams 31-34. The acquired one of the plurality of structured data records of each of the data streams 31-34 may also be referred to as a "sample" of the respective data stream 31-34 and/or as "semantics information", that is, information from which a semantics of the respective data stream 31-34 may be derived.

20 The registry device 5 may subscribe to each of the data streams 31-34 to receive a respective sample directly from a respective data provider device 21-24 and/or may be provided with a respective sample from the tagging device 8, which is subscribed to each of the data streams 31-34, as will be explained hereinbelow.

The registry device 5 then parses each of the acquired samples to determine a respective semantics specification 61-64 of the respective data stream 31-34.

30 More particularly, for each of the predetermined tags, a number of patterns may be defined or stored in the registry

device 5. For example, for the predetermined tag "<value>", a pattern such as "*valu*", a pattern such as "*temp*" and a pattern such as "*pres*" may be defined in the registry device 5. Likewise, for the predetermined tag "<name>" a pattern such as 5 as "*nam*|*id*" may be defined in the registry device 5.

The registry device 5 may match each of the patterns defined for each of the predetermined tags against the field identifiers comprised by each of the received samples.

For example, for the first data stream 61, upon receiving the 10 following sample,

```
sensorValue:  350.7
timestamp:    1528444828
id:           press1
```

the registry device 5 may determine that the field identifier 15 "sensorValue" matches the pattern "*valu*" defined for the predetermined tag "<value>" and may therefore add a mapping between the predetermined tag "<value>" and the field identifier "sensorValue" to the semantics specification 61 of the first data stream 31.

20 Similar pattern matching processing may be repeated for each field identifier of each of the samples and will not be described in detail.

In this way, the semantics specifications 61-64 stored in the registry device 5 may be determined automatically.

25 According to the present embodiment, the consumer devices 71, 72 do not communicate with the registry device 5 to obtain the semantics specifications 61-64. Rather, the tagging device 8 performs tagging of the data streams 31-34, which is now described in more detail.

30 The tagging device 8 may be subscribed to each of the data streams 31, 32, 33, 34.

More specifically, according to one example, the tagging device 8 receives the first data stream 31 provided by the first data provider device 21 and the second data stream 32 provided by the second data provider device 22. The third data stream 33 provided by the third data provider device 23 and the fourth data stream 34 provided by the fourth data provider device 24 are received by a further computing resource 13 of the streaming platform 1. The further computing resource 13 may perform processing such as preprocessing, sanity checking, and merging of the third and fourth data streams 33, 34 and may transmit a combined data stream 33, 34 to the tagging device 8.

Thus, the tagging device 8 receives the first and second data streams 31, 32 and the combined data stream 33, 34. The tagging device 8 further queries the registry device 5 for the corresponding semantics specifications 61-64 and receives the semantics specifications 61-64 from the registry device 5.

For each of the received data streams 31-34, the tagging device 8 creates a modified data stream 312, 332, 342 by inserting, based on the received semantics specifications 61-64, instances of the predetermined tags "<value>" and "<name>" into each of the plurality of structured data records (4 in Fig. 1).

More specifically, the tagging device 8 inserts a respective predetermined tag (an instance thereof) into a respective structured data record such that the inserted instance of the respective predetermined tag is positionally associated with (such as placed next to, before or after) a field identifier corresponding to the respective predetermined tag according to the mapping comprised by the respective semantics specification 61-64.

For example, if a structured data record of the first data stream 31 may be represented as follows,

```
sensorValue: 350.7
timestamp: 1528444828
id: press1
```

5 and if the semantics specification 61 of the first data stream
31 comprises a mapping that may be represented as follows:

```
<value> : sensorValue
<name> : id
```

10 then the tagging device 8 may, for example, create a
structured data record of the modified first data stream 312
that may be represented as follows:

```
sensorValue<value>: 350.7
timestamp: 1528444828
id<name>: press1
```

15 Likewise, if two structured data records of the combined data
stream 33, 34 may be represented as follows,

```
manufacturer: SIE
sensorName: temp1
temperature: 1321.8
20 utime: 1528444829
```

```
manufacturer: SIE
sensorName: temp2
temperature: 1139.3
25 utime: 1528444829
```

and if the semantics specifications 63, 64 of the third and
fourth data streams 33, 34 comprise a mapping that may be
represented as follows,

<value> : temperature

<name> : sensorName

then the tagging device 8 may, for example, create two structured data records of the modified combined data stream

5 332, 342 that may be represented as follows:

```
manufacturer:      SIE
sensorName<name>:   temp1
temperature<value>: 1321.8
utime:             1528444829
```

10

```
manufacturer:      SIE
sensorName<name>:   temp2
temperature<value>: 1139.3
utime:             1528444829
```

15 The tagging device 8 then forwards the modified first data stream 312 to the first data consumer device 71 and forwards the modified combined third and fourth data stream 332, 342 to the second data consumer device 72.

Similar tagging processing may also be performed for the
20 second data stream 32 and a further data consumer device (not shown) consuming the second data stream 32.

As can be appreciated from the exemplary representations of exemplary structured data records of the modified data streams 312, 332, 342 shown above, the data consumer devices 71, 72
25 may advantageously be able to access the pieces of measurement data of interest (i.e. the values 350.7, 1139.3, 1321.8 and the corresponding sensor names) based on the instances of the predetermined tags inserted into the respective modified data streams 312, 332, 342.

30

Specifically, when receiving a respective structured data record of the modified data stream 312, the consumer device 71 may iterate through each field of the structured data record. For each field, the consumer device 71 may determine whether a field identifier of the field is tagged with one of the predetermined tags "<value>" or "<name>". If the field identifier is tagged with the predetermined tag "<value>" (field identifier "sensorValue" in the example above), the consumer device 71 may access the value of the field (350.7) and identify the accessed value as the piece of measurement data of interest, i.e. as the sensor value. If the field identifier is tagged with the predetermined tag "<name>" (field identifier "id" in the example above), the consumer device 71 may access the value of the field and identify the accessed value ("press1") as a further piece of measurement data of interest, i.e. as the sensor name. If the field identifier is not tagged, or is tagged with an unknown tag, the consumer device 71 may disregard the value of the field and move on to the next field.

Similar processing is carried out by the data consumer device 72 and a description thereof will not be repeated for brevity.

The data consumer devices 71, 72 may consume the pieces of measurement data of interest obtained in such a way for any purpose related to, for example, analysis, monitoring and/or control of the industrial facility (10 in Fig. 1).

For example, according to the third exemplary embodiment, the data consumer device 71 may store the pieces of measurement data such obtained in a historical database 14 for later retrieval. Storing the pieces of measurement data in this way may require less storage space than storing the entire first data stream 31, since pieces of measurement data of no interest (such as the "manufacturer" field in the examples above) may be discarded.

For example, the second data consumer device 72 may provide the obtained measurement data to the external unit 11, which may be a third-party unit running a third-party software for analysis of the industrial facility (10 in Fig. 1).

5 According to another variant, for example, processing for analysis, monitoring and/or control of the industrial facility (10) based on the obtained measurement data may be performed by one or more of the data consumer devices 71, 72.

10 Although the present invention has been described in accordance with preferred embodiments, it is obvious for the person skilled in the art that modifications are possible in all embodiments.

The tagging device 8, the registry device 5 and a respective data consumer device 71, 72 have been described as separate, 15 single devices. However, one or more of the devices 8, 5, 71, 72 may also be embodied together as a single device, or may be embodied as a plurality of distributed interconnected computing resources arranged in a data center or cloud.

20 By implementing the streaming platform 1 in a data center or cloud, sufficient computing power may be provided to perform the processing of a large number of high-frequency data streams acquired with a frequency of 1 kHz or above in real time.

25 In the second exemplary embodiment, the semantics specifications 61-64 are received by the data consumer devices 71, 72 and the data consumer devices 71, 72 extract portions of the measurement data based on the semantics specifications 61-64. In the fourth embodiment, the semantics specifications 61-64 are provided to the tagging device 8, and the data 30 consumer devices 71, 72 extract portions of the measurement data based on instances of the predetermined tags inserted into the modified data streams 312, 332, 342 by the tagging

device 8. However, also the second exemplary embodiment may use the tagging device 8, and also in the third exemplary embodiment, the data consumer devices 71, 72 may directly receive the semantics specifications from the tagging device

5 5.

In the second exemplary embodiment, the registry device 5 provides an interface for registering the tagging specifications 61-64, whereas in the third exemplary embodiment, the registry device 5 employs pattern matching to automatically determine the semantics specifications 61-64.

10 However, also in the second exemplary embodiment, the registry device 5 may employ pattern matching, and also in the third exemplary embodiment, the registry device 5 may provide an interface for registering semantics specifications 61-64 and/or for verifying and modifying the semantics specifications 61-64 that were automatically determined.

The number of predetermined tags is not limited to the examples "<value>" and "<id>". If further pieces of the measurement information, such as a timestamp, and the like, are of interest, further predetermined tags, such as "<time>", may be defined in the proposed system 1.

20 The present invention may be identified as a data-center or cloud-based streaming platform 1 for providing, consuming, and performing processing based on a plurality of data streams 3, 31-34, wherein information on a respective semantics of a respective data stream 3, 31-34 is made available to consumer devices 7, 71, 72 in the form of semantics specifications 6, 61-64 stored in a registry device 5 and/or in the form of tags inserted by a tagging device 8 into the plurality of data streams 3, 31-34. A respective data stream 3, 31-34 may include measurement data acquired from a data source 9 in an industrial facility 10, and the processing performed by the

streaming platform 1 may be related to monitoring, analysis, and control of the industrial facility 10.

Specifically, no specific configuration, interface component and the like may be required on the side of the data provider devices 2, 21-24 providing the plurality of data streams 3, 31-34. More particularly, a respective data provider device 2, 21-24 may provide a respective data stream 3, 31-34 using a respective native semantics of the respective data provider device 2, 21-24. Likewise, no specific configuration, interface component or knowledge about the various native semantics used by the respective data provider devices 2, 21-24 may be required on the side of the data consumer devices 7, 71, 72.

REFERENCE NUMERALS:

	1	system for streaming measurement data
	2, 21-24	provider device
5	3, 31-34	data stream
	4	structured data records
	5	registry device
	6, 61-64	semantics specification
	7, 71, 72	consumer device
10	8	tagging device
	9	data source
	10	industrial facility
	11	external unit
	12	assistance device
15	13	further computing resource
	14	historical database
	312, 332, 342	modified data stream

Patent claims

1. A system (1) for streaming measurement data comprising:

a number of provider devices (2, 21-24) each configured to provide one or more of a number of data streams (3, 31-34) each including a plurality of structured data records (4) indicative of measurement data obtained from one of a number of data sources (9) of an industrial facility (10);

a registry device (5) configured to store, for each of the number of data streams (3, 31-34), a semantics specification (6, 61-64) indicative of a semantics of the plurality of structured data records (4) of the respective data stream (3, 31-34); and

a number of consumer devices (7, 71, 72) configured to, for at least one of the number of data streams (3, 31-34):

- receive the respective data stream (3, 31-34) provided by the corresponding provider device (2, 21-24);

- obtain the measurement data from the plurality of structured data records (4) included in the received data stream (3, 31-34) based on the semantics specification of the received data stream (3, 31-34) stored in the registry device (5); and

- consume the obtained measurement data of the received data stream (3, 31-34).

2. The system according to claim 1, wherein the registry device (5) is configured to provide an interface for registering one or more of the semantics specifications (6, 61-64).

3. The system according to claim 1 or 2, wherein the registry device (5) is configured to determine one or more of the semantics specifications (6, 61-64) by parsing semantics

information about the corresponding data stream (3, 31-34), said semantics information being received from the corresponding provider device (2, 21-24).

4. The system according to claim 3, wherein the semantics information received by the registry device (5) comprises at least one of the plurality of structured data records (4) of the corresponding data stream (3, 31-34).

5. The system according to any of claims 1 to 4, wherein

a respective of the plurality of structured data records (4) comprises a number of fields, each field including a value and a field identifier associated with the value, and

a respective of the number of semantics specifications (6, 61-64) comprises a mapping between each of a number of predetermined tags and a corresponding one of the field identifiers.

6. The system according to claim 5, wherein said parsing of the received semantics information includes matching each of a number of patterns defined for each of the number of predetermined tags against each of a number of field identifiers included in the received semantic information.

7. The system according to claim 5 or 6, wherein at least one of the number of consumer devices (7, 71, 72) is further configured to receive the semantics specification (6, 61-64) of the respective data stream (3, 31-34) from the registry device (5).

8. The system according to claim 5 or 6, further comprising a tagging device (8) configured to, for at least one of the number of data streams (3, 31-34):

- receive, from the registry device (5), the semantics specification of the respective data stream (3, 31-34),

- receive the respective data stream (3, 31-34) provided by a corresponding one of the provider devices (2, 21-24),

- modify the received data stream (31, 32, 33, 34) by inserting, based on the received semantics' specification, instances of one or more of the number of predetermined tags into the plurality of structured data records of the received data stream (3, 31-34) to obtain a modified data stream (312, 332, 342); and

- forward the modified data stream (312, 332, 342) to each of the number of consumer devices (71, 72) consuming the corresponding data stream (31, 33, 34).

9. The system according to claim 8, wherein the tagging device (8) is configured to, for each field of each of the plurality of structured data records, if it is determined, based on the semantics specification, that the field identifier of the respective field corresponds to one of the predetermined tags, associate an instance of the corresponding one of the predetermined tags with the field identifier.

10. The system according to any of claims 5 to 9, wherein a respective one of the number of consumer devices (7, 71, 72) is configured to obtain the measurement data from the plurality of structured data records of the at least one of the number of data streams (3, 31-34) by accessing a number of fields of a respective structured data record, each of the accessed number of fields being a field having an associated field identifier that is associated with one of the predetermined tags.

11. The system according to any of claims 1 to 10, wherein consuming the obtained measurement data includes one or more of:

providing the obtained measurement data to an external unit connected to the system (1) for streaming measurement data;

storing the obtained measurement data into a historical database (13) comprised by the system (1) for streaming measurement data; and/or

performing, on the consumer device (7, 71, 72) and/or a number of further devices connected to the consumer device and comprised by the system for streaming measurement data (1), at least one of analysis, monitoring, and control of the industrial facility (10) using the obtained measurement data.

12. The system according to any of claims 1 to 11, wherein the measurement data is high-frequency measurement data obtained from the number of data sources (9) with a frequency of 1 kHz or higher.

13. The system according to any of claims 1 to 12, wherein the number of data streams comprise a number of real-time data streams provided by the number of provider devices and received by the number of consumer devices in real time.

14. A method for streaming measurement data comprising:

providing (S1), by each of a number of provider devices (2, 21-24), one or more of a number of data streams (3, 31-34) each including a plurality of structured data records (4) indicative of measurement data obtained from one of a number of data sources (9) of an industrial facility (10);

storing (S2), in a registry device (5), for each of the number of data streams (3, 31-34), a semantics specification (6, 61-4) indicative of a semantics of the plurality of structured data records (4) of the respective data stream (3, 31-34); and

performing, by a number of consumer devices (7, 71, 72) and for at least one of the number of data streams (3, 31-34), the steps of:

- receiving (S3) the respective data stream (3, 31-34) provided by the corresponding provider device (2, 21-24);

- obtaining (S4) the measurement data from the plurality of structured data records (4) included in the received data stream (3, 31-34) based on the semantics specification of the received data stream (3, 31-34) stored in the registry device

5 (5); and

- consuming (S4) the obtained high-frequency measurement data of the received data stream (3, 31-34).

15. A computer program product comprising a program code for executing the method of claim 14 when run on at least one

10 computer.

FIG 1

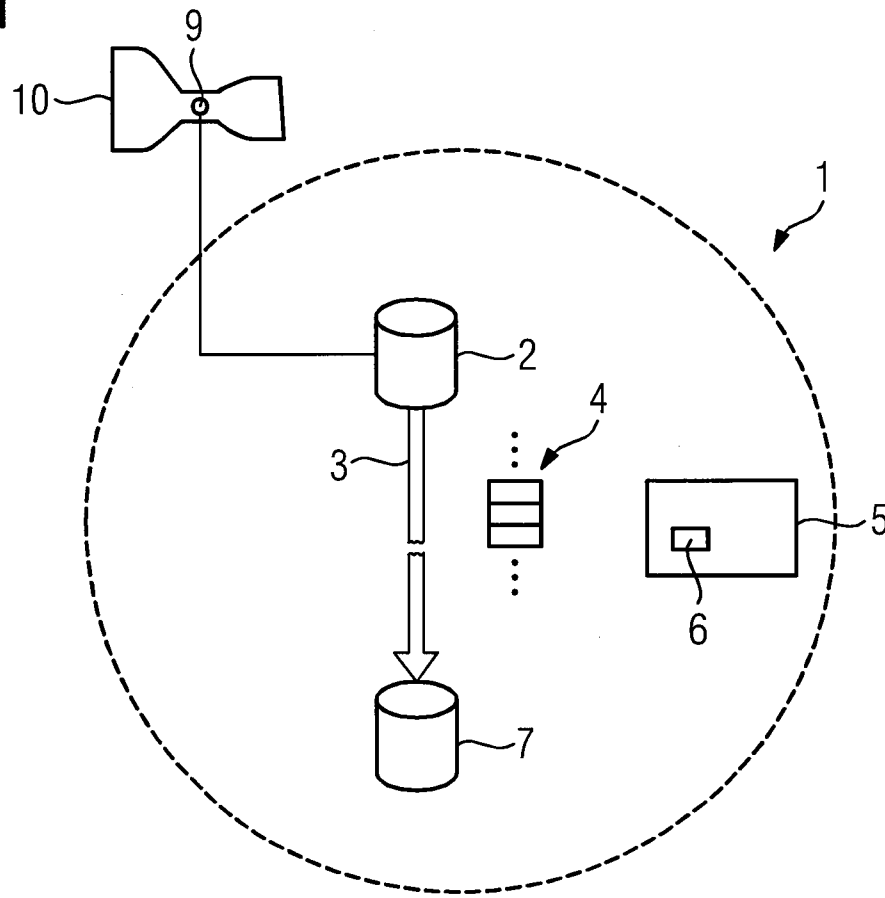


FIG 2

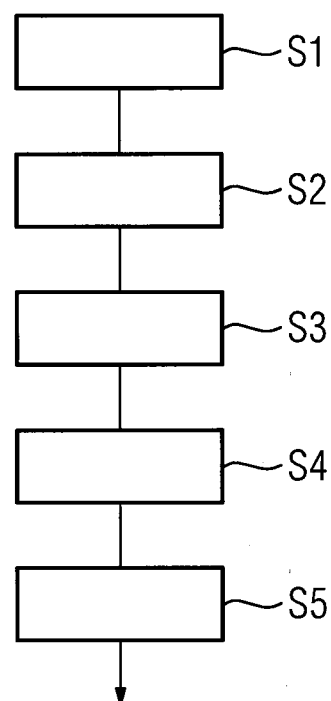


FIG 3

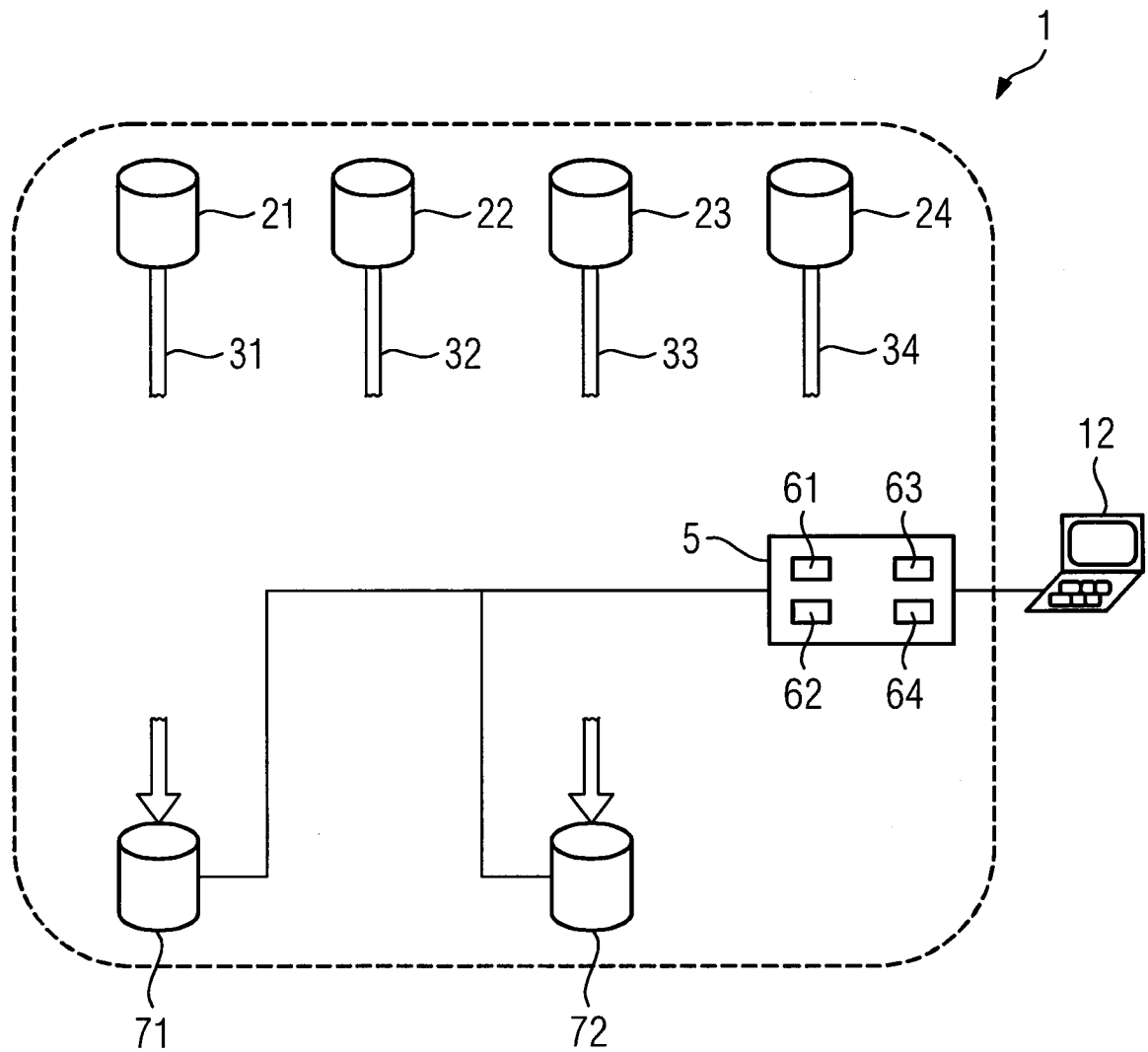
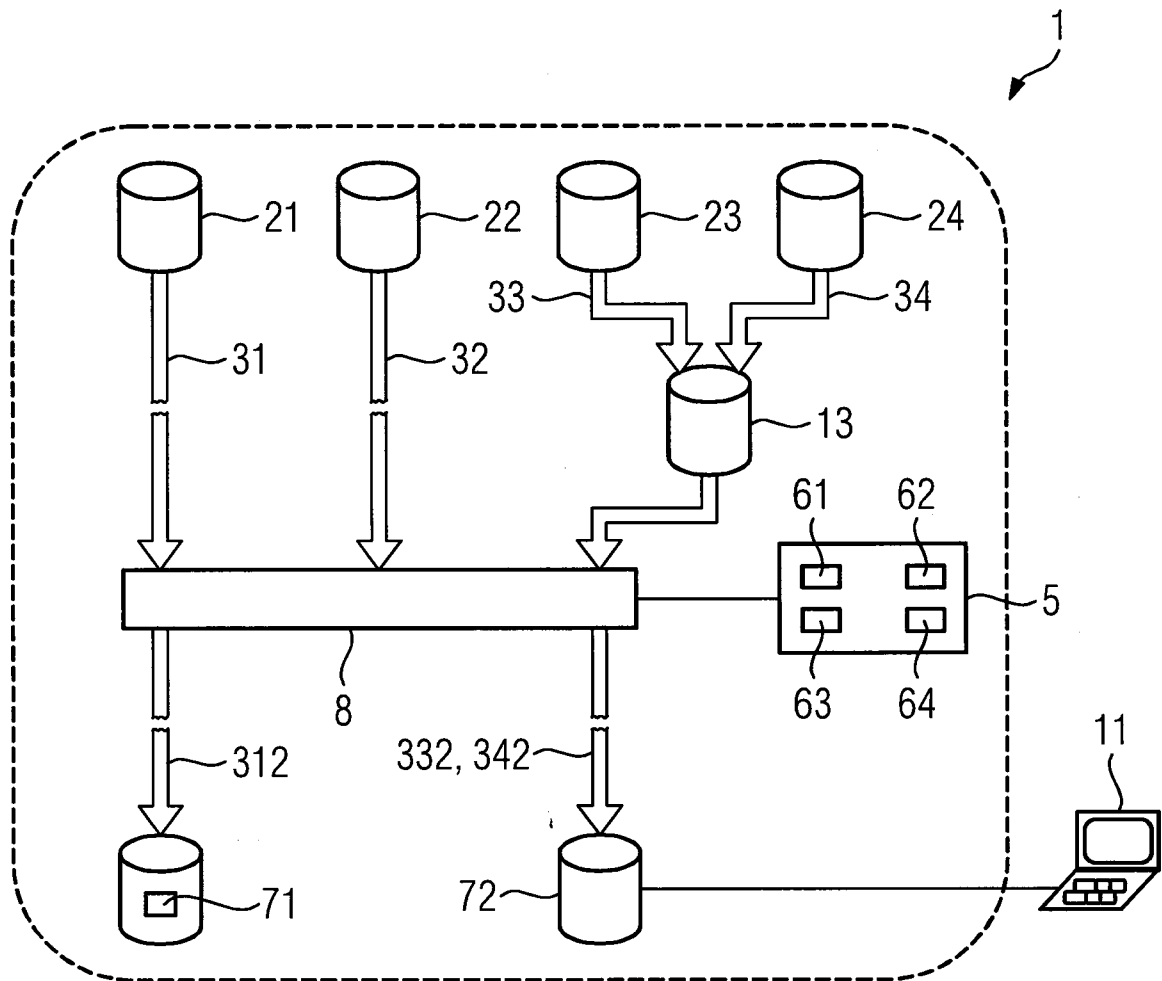


FIG 4



INTERNATIONAL SEARCH REPORT

International application No
PCT/RU2018/000420

A. CLASSIFICATION OF SUBJECT MATTER
INV. G05B19/418
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)
H04L G05B

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 2016/315816 A1 (FANG JOSEPH Y [US]) 27 October 2016 (2016-10-27) paragraph [0008] - paragraph [0084]; figures 1-7	1-15
X	US 2017/124166 A1 (THOMAS LOUIS KELLY [US] ET AL) 4 May 2017 (2017-05-04) paragraph [0002] - paragraph [0009]; figures 1-12,15 paragraph [0186] - paragraph [0192] paragraph [0213] - paragraph [0225]	1-15



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

6 March 2019

Date of mailing of the international search report

15/03/2019

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

Barriuso Poy, Alex

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/RU2018/000420

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016315816	A1	27-10-2016	NONE

US 2017124166	A1	04-05-2017	NONE
