(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 August 2020 (20.08.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/167464 A1

(51) International Patent Classification:
*G06F 11/00* (2006.01) *G06F 11/34* (2006.01)
*G06F 11/07* (2006.01) *G06F 11/32* (2006.01)

(21) International Application Number:
PCT/US2020/015508

(22) International Filing Date:
29 January 2020 (29.01.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/804,639 12 February 2019 (12.02.2019) US
16/416,107 17 May 2019 (17.05.2019) US

(71) Applicant: **MICROSOFT TECHNOLOGY LICENSING, LLC** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: **DARWIN, Cosmos**; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **MATTHEW, Bryan S.**; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **MACGREGOR, Donald**; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **LEE, Scott Chao-Chueh**; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(74) Agent: **MINHAS, Sandip S.** et al.; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(54) Title: FAULT PREDICTION AND DETECTION USING TIME-BASED DISTRIBUTED DATA

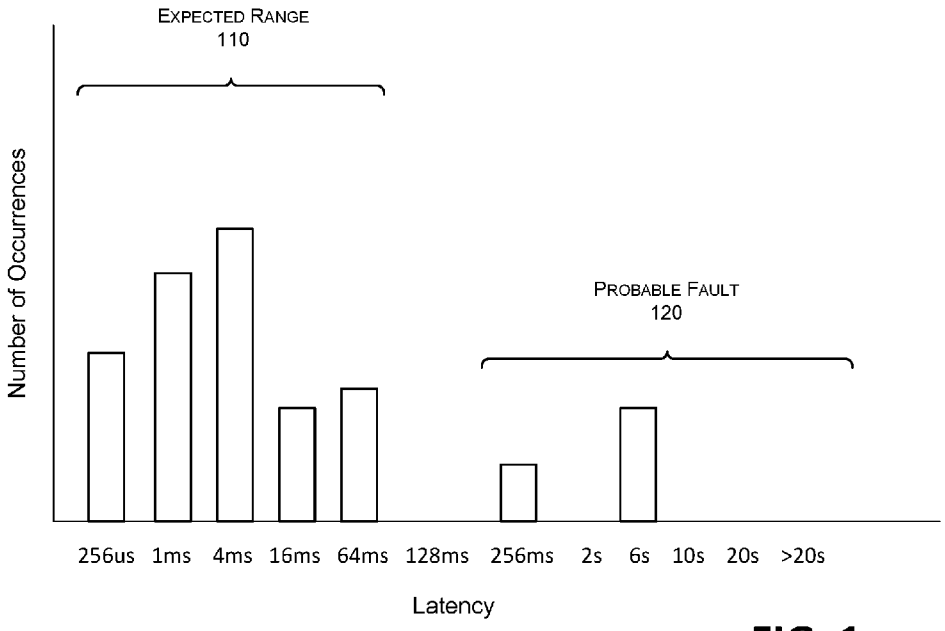



FIG. 1

(57) Abstract: Performance data is collected for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network. Based on the collected performance data, a time-based I/O performance profile for the storage device is determined. A characteristic time-based I/O performance profile is determined for a representative group of storage devices having common characteristics with the storage device and based on previously collected performance data for devices of the representative group. It is determined that the difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of a probable failure of the storage device. An indication is generated that the storage device exceeded the predetermined deviance threshold.

WO 2020/167464 A1

[Continued on next page]

(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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, 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).

**Declarations under Rule 4.17:**

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**

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

# FAULT PREDICTION AND DETECTION USING TIME-BASED DISTRIBUTED DATA

## BACKGROUND

**[0001]** A data center is a facility that houses computer systems and various networking, storage, and other related components. Data centers may, for example, provide computing services to businesses and individuals as a remote computing service or provide "software as a service" (e.g., cloud computing).

**[0002]** When a failure of a component such as a storage device occurs, data center users may lose data and may be unable to provide services to their downstream clients, resulting in lost revenue and user dissatisfaction. Troubleshooting storage and other I/O issues may be difficult to perform given the complexity of the networks in data centers. Production loss and inefficiencies with respect to computing resources can be exacerbated when the data center is unable to quickly isolate and correct the cause of a device issue. It is with respect to these considerations and others that the disclosure made herein is presented.

## SUMMARY

**[0003]** It is desirable to provide the highest level of data storage availability and resiliency while at the same time providing performance and minimizing cost. It is also desirable to avoid faults and to take corrective or preemptive action before faults occur. In a system with storage devices, a performance issue in a single storage device can lead to performance issues for storage resources that are allocated across multiple storage devices, for example when mirroring or parity resiliency schemes are implemented. As used herein, such resiliency schemes may also be referred to as a resiliency function or resiliency mechanism.

**[0004]** Some methods attempt to identify failures using specialized and disruptive diagnostics such as installing special firmware or injecting added data traffic, which can cause delays or disruptions and may produce mixed results. Some storage systems attempt to identify faults by measuring the average latency for input/output (I/O) operations to individual storage devices. However, taking the average may mask many performance issues. For example, existing techniques that measure the average latency do not readily identify device failure modes that involve a relatively small number of I/O operations that exhibit excessive latency.

**[0005]** The disclosed embodiments describe technologies for generating and utilizing a distribution of I/O latency and other performance characteristics to identify potential device failures in a system that includes storage devices, such as a distributed

software-defined storage system. The distribution may be based on time or other bases. For example, one distribution can be I/O latency values plotted on a time axis. Tail latencies, or higher percentile latencies, may be used. This can provide advantages to resiliency/redundancy schemes, where storage devices are combined into logical units using mirroring or parity or other erasure coding schemes. Taking a storage device offline or allowing the storage device to fail before replacing the device may significantly impact the resiliency guarantees associated with such schemes. For example, a storage device that is exhibiting behavior that is indicative of a potential failure can be load balanced with some I/O requests being routed to other devices. By diverting some requests, the device may be allowed to internally recover and eventually resume normal operation. Alternatively, a replacement device may be selected and populated with data so that it can replace the potentially failing device before such a failure actually occurs. Other preemptive actions may also be taken.

**[0006]** In one aspect, a user interface may be implemented that allows users to request and selectively view I/O latency and other data over a configurable time-based or histogram-based distribution. The user interface enables comparison of the I/O latency distribution to data from other devices in the same or similar class to identify potential failures. This may allow data centers to more effectively adhere to operational objectives and at the same time improve operating efficiencies.

**[0007]** In many data centers, I/O issues may cause patterns of delays or drops of a percentage of data packets. I/O issues may also affect data storage resiliency. For example, physical disk drive components may be combined into one or more logical units to provide data redundancy and performance improvement. Data may also be distributed across the drives depending on the desired level of redundancy, performance, reliability, availability, and capacity. Different levels of resiliency can be achieved, for example, by different mirroring schemes or parity schemes.

**[0008]** Issues with storage devices can have a negative impact on performance and the user experience and may be difficult to isolate. As the data center scale grows, detecting and fixing these faulty devices are becoming even more important. At the same time, identifying such devices can be difficult, and unfocused attempts to collect latency data may result in a large amount of data that must be processed potentially without generating any actionable intelligence. The data center can thus spend a considerable amount of time and effort attempting to isolate I/O faults, which can lead to extensive and unnecessary consumption of computing resources.

**[0009]** In a system with storage devices, a performance issue in a single storage device can lead to performance issues for storage resources that are allocated across multiple storage devices. Some methods attempt to identify failures using specialized and disruptive diagnostics such as installing special firmware or injecting added data traffic, which can cause delays or disruptions and may produce mixed results. Some storage systems attempt to identify faults by measuring the average latency for input/output (I/O) operations to individual storage devices. However, taking the average may mask many performance issues. For example, existing techniques that measure the average latency do not readily identify device failure modes that involve a relatively small number of I/O operations that exhibit excessive latency. If such issues are not identified in a timely manner, downtime and data loss may occur, leading to a poor user experience.

**[0010]** In various embodiments, methods and systems are disclosed for generating and utilizing a distribution of I/O latency and other characteristics to identify potential device failures and states in a storage network. A user interface may be implemented that allows users to request and selectively view the I/O latency and other data over a configurable time-based or histogram-based distribution. The user interface further enables comparison of the I/O latency distribution to data from other devices in the same class to identify potential failures. For example, the UI may provide a histogram of latency ranges and the number of I/O operations with latencies that fell within each range. The UI can be rendered on a display or implemented as a programming interface where the data can be returned to a requesting client.

**[0011]** In some embodiments, the UI can also be configured to specify threshold values for I/O latencies and view a time-based distribution. For example, a system can receive, via the UI, a latency threshold and a time range from a user input. The latency threshold and time range can be utilized to control the display of I/O operations that exceed the threshold, and their distribution over time.

**[0012]** In some embodiments, the system can further be configured to provide notifications that may be triggered when a percentage of I/O operations exceed a specified latency for a given time duration.

**[0013]** In some embodiments, techniques may be implemented for using the distributed I/O latency and other time-based data to predict storage device failures and other device states. The time-based distributed data can be referred to herein as distributed behavioral and characteristic data. The distributed behavioral and characteristic data can be used to identify outlying or deviant behavior as compared to other devices within the same

class or grouping. For example, a device with a latency profile that is statistically deviant compared to a characteristic profile for other devices of the same class or grouping can indicate a high probability that a failure will occur within a threshold period of time. The system can automatically identify and report potential failures based on thresholds and parameters that are configured via the UI. The failure probability can be adjusted based on actual observed failures.

**[0014]** In an embodiment, the threshold can be predetermined. The threshold may also be dynamically updated. The threshold may be based on empirical observations of I/O data for a group of devices. The threshold can also be based on one or more statistical factors such as standard deviation. For example, the threshold can be three or five times the standard deviation for a given parameter.

**[0015]** The behavioral and characteristic data can include the time distribution of I/O latencies as well as other data that can be readily observable or self-reported. Such data can include failure and maintenance data, device temperature, and I/O queue time. The behavioral and characteristic data can be collected across entire populations of devices at the same or across data centers. In one embodiment, the behavioral and characteristic data can exclude the data for the current device of interest. Characteristic patterns can then be identified that are indicative of failures for various classes of devices, allowing for preemptive device maintenance that can be used to avoid downtime and data loss.

**[0016]** This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended that this Summary be used to limit the scope of the claimed subject matter. Furthermore, the claimed subject matter is not limited to implementations that solve any or all disadvantages noted in any part of this disclosure.

## DRAWINGS

**[0017]** The Detailed Description is described with reference to the accompanying figures. In the description detailed herein, references are made to the accompanying drawings that form a part hereof, and that show, by way of illustration, specific embodiments or examples. The drawings herein are not drawn to scale. Like numerals represent like elements throughout the several figures.

**[0018]** FIG. 1 is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0019]** FIG. 2A is a diagram illustrating an example user interface in accordance

with the present disclosure;

**[0020]** FIG. 2B is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0021]** FIG. 2C is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0022]** FIG. 2D is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0023]** FIG. 2E is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0024]** FIG. 2F is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0025]** FIG. 3A is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0026]** FIG. 3B is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0027]** FIG. 3C is a diagram illustrating an example user interface in accordance with the present disclosure;

**[0028]** FIG. 4 is a diagram illustrating an example architecture in accordance with the present disclosure;

**[0029]** FIG. 5 is a diagram illustrating an example architecture in accordance with the present disclosure;

**[0030]** FIG. 6 is a diagram illustrating an example architecture in accordance with the present disclosure;

**[0031]** FIG. 7 is a diagram illustrating an example architecture in accordance with the present disclosure;

**[0032]** FIG. 8 is a flowchart depicting an example procedure for diagnosing storage issues in accordance with the present disclosure;

**[0033]** FIG. 9 is a flowchart depicting an example procedure for diagnosing storage issues in accordance with the present disclosure;

**[0034]** FIG. 10 is an example computing device in accordance with the present disclosure.

## DETAILED DESCRIPTION

**[0035]** The following Detailed Description describes technologies for diagnosing storage issues in data centers and other networked environments, including those that

provide virtualized services. In particular, disclosed are techniques for generating and utilizing the distribution of I/O latency and other performance characteristics to identify potential device failures in a system that includes storage devices, such as a distributed software-defined storage system. The storage systems may include data backup and recovery schemes such as mirroring and parity schemes. The improvements can include improved data resiliency, reduced downtime, and reduced costs.

**[0036]** The distribution can be over time or other unit that allows for analysis of I/O-related parameters to be analyzed in ways other than using a time-based average. In one aspect, a user interface is implemented that allows users to request and selectively view I/O latency and other data over a configurable time-based or histogram-based distribution. The user interface further enables comparison of the I/O latency distribution to data from other devices in the same class to identify potential failures.

**[0037]** In a system with storage devices, a performance issue in a single storage device can lead to performance issues for storage resources and services that are allocated across multiple storage devices. Some methods attempt to identify failures using specialized and disruptive diagnostics such as installing special firmware or injecting added data traffic, which can cause delays or disruptions and may produce mixed results. Some storage systems attempt to identify faults by measuring the average latency for input/output (I/O) operations to individual storage devices. However, taking the average may mask many performance issues. For example, existing techniques that measure the average latency do not readily identify device failure modes that involve a relatively small number of I/O operations that exhibit excessive latency. Device self-reporting of fault messages may be used, after a fault occurs. However, these methods either fail to identify issues that can be detected using the distribution of data, or fail to identify issues before they manifest into performance issues. It would be useful to identify and predict performance issues and failures of a device, before the failures and performance issues actually occur. This can allow the service provider to take preemptive measures, such as bringing another device on-line, and thus avoiding problems associated with an unexpected loss of a device. If such issues are not identified in a timely manner, downtime and data loss may occur, leading to a poor user experience.

**[0038]** The present disclosure describes ways to generate and utilize the distribution of I/O latency and other characteristics to identify potential device failures and states in a network that utilizes storage devices. In one embodiment, failures that have a threshold probability of occurring are determined so that the service provider can replace devices before a failure actually occurs. The failure predictions may be determined using historical

and real time data for the storage devices. The service provider may determine criteria such as a probability at which corrective action should be taken, such as replacing the device, or reducing workload at the device. For example, the provider network can identify a disk drive as having a probability of failure of 95% over the next 10 days. The service provider can determine an objective that it wishes to achieve, such as to have zero actual failures for live devices with a given confidence level, or minimization of I/O latencies that exceed a threshold time period.

**[0039]** In one embodiment, a user interface may be implemented that allows users to request and selectively view the I/O latency and other data over a configurable time-based or histogram-based distribution. The user interface may be configured to enable comparison of the I/O latency distribution for a selected device to I/O latency distribution data for other devices in the same class to identify potential performance issues. For example, the UI may provide a histogram of latency ranges and the number of I/O operations with latencies that fell within each range. FIG. 1 illustrates one example of a UI in accordance with the present disclosure. In the example, a distribution of latencies for a selected storage device is depicted. The vertical axis shows the number of occurrences, and the horizontal axis shows the distribution of recorded latencies. In the figure, each bar represents the total number of I/O operations where the latency was greater than 256us, 1ms, 4ms, etc. The UI can be rendered on a display or implemented as a programming interface where the data can be returned to a requesting client for rendering on a client device. FIG. 1 illustrates a normal range 110 that indicates I/O responses that are considered within the normal range of response times for operations during a typical operational environment. The figure also illustrates a probable fault range 120 that indicates I/O responses that are considered within a range that is predictive of response times for operations that are marginal and may indicate a potential fault.

**[0040]** The UI can also be configured to specify threshold values for I/O latencies and view a time-based distribution. For example, a system can receive, via the UI, a latency threshold and a time range based on a user input. The latency threshold and time range can be utilized to control the display of I/O operations that exceed the threshold, and their distribution over time. FIG. 2 illustrates one example of a UI in accordance with the present disclosure, where latency values that exceeded 2 seconds are shown over a two-week period. In the example illustrated in FIG. 2, a time-based distribution of latencies for a selected storage device that have exceeded a specified latency threshold is depicted. The vertical axis shows the number of occurrences of latencies that have exceeded 2 seconds, and the

horizontal axis shows the distribution of recorded latencies over a specified time period. FIG. 2A illustrates an expected distribution 210 that indicates I/O responses that are considered within the normal range where some number of higher latencies may have occurred during a peak usage time for operations during a typical operational environment but does not otherwise indicate a problem. The figure also illustrates a marginal distribution 220 that indicates a larger number of I/O responses that occurred within a time period that are considered within a range that is predictive of pre-marginal or marginal behavior and indicates a probable fault.

**[0041]** In some embodiments, a UI may be configured to provide interactive capability through the use of an input device such as a mouse or touchscreen. In the example shown in FIG. 2B, a user may select 230 a set of occurrences of interest. FIG. 2C illustrates that in response to the user selection, further information 240 may be provided for the selected set of occurrences. In one example, the information may include date, time, and number of occurrences.

**[0042]** FIG. 2D illustrates another example where a user may select a range 250 of occurrences for interaction. FIG. 2E illustrates that in response to the user selection, further information 260 may be provided for the selected group of occurrences. In the example, the user may be provided options for accessing more information, viewing profile and other information for similar devices in the same group or elsewhere, set a threshold for further monitoring, and access further actions.

**[0043]** FIG. 2F illustrates that in response to a further user selection for more information, further information 270 may be provided for the selected group of occurrences. In one example, the information may include date, time, and number of occurrences for each of the histograms in the selected group.

**[0044]** In an embodiment, the UI and/or a notification function can be configured to provide notifications that may be triggered when a percentage of I/O operations exceed a specified latency for a given time duration. FIG. 3A illustrates one example of a UI in accordance with the present disclosure, where an alert is generated under specified conditions. FIG. 3A illustrates an example of a user interface 300 where storage management interface 310 is rendered. The interface 310 allows a user to specify a percentage of I/O operations 320, duration 330, and condition duration 340. The user may also specify whether the I/O data to be accessed are reads, writes, or a combination 350. The UI may return a depiction of the requested data 360.

**[0045]** FIG. 3B illustrates a further example where the interface 310 includes a

summary window 370 indicating that for the selected user criteria, there are no failed I/O's, and that there are approximately 1K slow I/O's that make up 0.02% of the I/O's for the device for the selected time period.

**[0046]** In another example, FIG. 3C illustrates a further example where the interface 310 includes a warning window 380 indicating that for the selected device, a difference in the profile has been detected when compared to performance profiles for other devices, and that the difference may indicate that the drive is malfunctioning, damaged, or beginning to fail.

**[0047]** The disclosure further describes techniques for using the distributed I/O latency and other time-based distributed data to predict storage device failures and other device states. The time-based distributed data can be referred to generally as behavioral and characteristic data. In some embodiments, the behavioral and characteristic data may be distributed over different units. For example, the behavioral and characteristic data may be plotted on a frequency axis.

**[0048]** In an embodiment, the distribution of I/O latency and other performance characteristics for a given storage device may be used to identify potential device issues. The distribution of I/O latency may be based on time or other bases. For example, tail latencies, or higher percentile latencies, may be used to identify potential device issues. A comparison may be performed of the I/O latency distribution for a storage device to data from other storage devices in the same class to identify whether the storage device is entering a pre-marginal or marginal state. The distribution of I/O latency and other performance characteristics may be used to determine a performance profile for storage device. When the performance profile is determined based on I/O performance information, then the performance profile may be referred to as a time-based I/O performance profile.

**[0049]** The time-based distributed data can be referred to herein as distributed behavioral and characteristic data. In an embodiment, the behavioral and characteristic may be used to identify outlying or deviant behavior as compared to other devices within the same class or grouping. For example, such devices may be considered in the same class or grouping based on being made by the same manufacturer, having similar technical specifications such as storage capacity and I/O speeds, built within a specified time period, and the like. For example, a device with a latency profile that is statistically deviant compared to a characteristic profile for other devices of the same class or grouping can indicate a high probability that a failure will occur within a threshold period of time. The system can automatically identify and report potential failures based on thresholds and

parameters that are configured via the UI. The failure probability can be adjusted based on actual observed failures. For example, if a machine learning function is used to learn and identify pre-fault behavior, then the behaviors of a failed device before the device failed may be used to further improve the training of the machine learning function.

**[0050]** The behavioral and characteristic data can include the time distribution of I/O latencies as well as other data that can be readily observable or self-reported. Such data can include failure and maintenance data, device temperature, and I/O queue time. The behavioral and characteristic data can be collected across entire populations of devices at the same or across data centers. Characteristic patterns can then be identified that are indicative of failures for various classes of devices, allowing for preemptive device maintenance that can be used to avoid downtime and data loss.

**[0051]** In one embodiment, the characteristic profiles may be developed until they are relatively deterministic with respect to a performance issue for a given parameter and for a given class or type of device. The distribution of I/O latency and other performance characteristics may be used to determine a characteristic performance profile for the given class or type of device. When the characteristic performance profile is determined based on I/O performance information, then the characteristic performance profile may be referred to as a characteristic time-based I/O performance profile.

**[0052]** The measurements may be analyzed using various methods, such as standard deviation or a time-based graph. By comparing the profile of a given device with the characteristic profiles, a status of the device may be determined, such as determining when and if the resource will enter into a faulty performance state. In some embodiments, a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile may be determined. The difference may be referred to as a deviance or a performance deviance. When the difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a threshold, then the storage device may be considered as being in a pre-marginal state. This threshold may be referred to as a deviance threshold or a predetermined deviance threshold. The threshold may be determined based on an analysis of historical data when devices have previously entered the pre-marginal state. The threshold may further be determined based on heuristics as well as machine learning. The threshold may be a single value, a range of values, or may be represented in other ways such as a table, graph, or a function.

**[0053]** In some embodiments, the threshold may be determined based on a

difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile at a plurality of time periods.

**[0054]** FIG. 4 illustrates an example computing environment in which the embodiments described herein may be implemented. FIG. 4 illustrates a data center 400 that is configured to provide computing resources to users 400a, 400b, or 400c (which may be referred herein singularly as "a user 400" or in the plural as "the users 400") via user computers 402a,402b, and 402c (which may be referred herein singularly as "a computer 402" or in the plural as "the computers 402") via a communications network 430. The computing resources provided by the data center 400 may include various types of resources, such as computing resources, data storage resources, data communication resources, and the like. Each type of computing resource may be general-purpose or may be available in a number of specific configurations. For example, computing resources may be available as virtual machines. The virtual machines may be configured to execute applications, including Web servers, application servers, media servers, database servers, and the like. Data storage resources may include file storage devices, block storage devices, and the like. Each type or configuration of computing resource may be available in different configurations, such as the number of processors, and size of memory and/or storage capacity. The resources may in some embodiments be offered to clients in units referred to as instances, such as virtual machine instances or storage instances. A virtual computing instance may be referred to as a virtual machine and may, for example, comprise one or more servers with a specified computational capacity (which may be specified by indicating the type and number of CPUs, the main memory size and so on) and a specified software stack (e.g., a particular version of an operating system, which may in turn run on top of a hypervisor).

**[0055]** Data center 400 may include servers 416a, 416b, and 416c (which may be referred to herein singularly as "a server 416" or in the plural as "the servers 416") that provide computing resources available as virtual machines 418a and 418b (which may be referred to herein singularly as "a virtual machine 418" or in the plural as "the virtual machines 418"). The virtual machines 418 may be configured to execute applications such as Web servers, application servers, media servers, database servers, and the like. Other resources that may be provided include data storage resources (not shown on FIG. 1) and may include file storage devices, block storage devices, and the like. Servers 416 may also execute functions that manage and control allocation of resources in the data center, such as a controller 415. Controller 415 may be a fabric controller or another type of program

configured to manage the allocation of virtual machines on servers 416.

**[0056]** Referring to FIG. 4, communications network 430 may, for example, be a publicly accessible network of linked networks and may be operated by various entities, such as the Internet. In other embodiments, communications network 430 may be a private network, such as a corporate network that is wholly or partially inaccessible to the public.

**[0057]** Communications network 430 may provide access to computers 402. Computers 402 may be computers utilized by users 400. Computer 402a,402b or 402c may be a server, a desktop or laptop personal computer, a tablet computer, a smartphone, a set-top box, or any other computing device capable of accessing data center 400. User computer 402a or 402b may connect directly to the Internet (e.g., via a cable modem). User computer 402c may be internal to the data center 400 and may connect directly to the resources in the data center 400 via internal networks. Although only three user computers 402a,402b, and 402c are depicted, it should be appreciated that there may be multiple user computers.

**[0058]** Computers 402 may also be utilized to configure aspects of the computing resources provided by data center 400. For example, data center 400 may provide a Web interface through which aspects of its operation may be configured through the use of a Web browser application program executing on user computer 402. Alternatively, a stand-alone application program executing on user computer 402 may be used to access an application programming interface (API) exposed by data center 400 for performing the configuration operations.

**[0059]** Servers 416 may be configured to provide the computing resources described above. One or more of the servers 416 may be configured to execute a manager 420a or 420b (which may be referred herein singularly as "a manager 420" or in the plural as "the managers 420") configured to execute the virtual machines. The managers 420 may be a virtual machine monitor (VMM), fabric controller, or another type of program configured to enable the execution of virtual machines 418 on servers 416, for example.

**[0060]** It should be appreciated that although the embodiments disclosed above are discussed in the context of virtualized storage services, other types of implementations can be utilized with the concepts and technologies disclosed herein.

**[0061]** In the example data center 400 shown in FIG. 4, a network device 411 may be utilized to interconnect the servers 416a and 416b. Network device 411 may comprise one or more switches, routers, or other network devices. Network device 411 may also be connected to gateway 440, which is connected to communications network 430. Network device 411 may facilitate communications within networks in data center 400, for example,

by forwarding packets or other data communications as appropriate based on characteristics of such communications (e.g., header information including source and/or destination addresses, protocol identifiers, etc.) and/or the characteristics of the private network (e.g., routes based on network topology, etc.). It will be appreciated that, for the sake of simplicity, various aspects of the computing systems and other devices of this example are illustrated without showing certain conventional details. Additional computing systems and other devices may be interconnected in other embodiments and may be interconnected in different ways.

**[0062]** It should be appreciated that the network topology illustrated in FIG. 4 has been greatly simplified and that many more networks and networking devices may be utilized to interconnect the various computing systems disclosed herein. These network topologies and devices should be apparent to those skilled in the art.

**[0063]** It should also be appreciated that data center 400 described in FIG. 4 is merely illustrative and that other implementations might be utilized. Additionally, it should be appreciated that the functionality disclosed herein might be implemented in software, hardware or a combination of software and hardware. Other implementations should be apparent to those skilled in the art. It should also be appreciated that a server, gateway, or other computing device may comprise any combination of hardware or software that can interact and perform the described types of functionality, including without limitation desktop or other computers, database servers, network storage devices and other network devices, PDAs, tablets, smartphone, Internet appliances, television-based systems (e.g., using set top boxes and/or personal/digital video recorders), and various other consumer products that include appropriate communication capabilities. In addition, the functionality provided by the illustrated modules may in some embodiments be combined in fewer modules or distributed in additional modules. Similarly, in some embodiments the functionality of some of the illustrated modules may not be provided and/or other additional functionality may be available.

**[0064]** FIG. 5 is a diagram illustrating a storage analysis function 500 for identifying issues in accordance with the present disclosure. In FIG. 5, a storage analysis function 500 may include resources and applications that may execute, for example, on one or more servers 530. It will be appreciated that some embodiments may involve additional computing resources of various types that may be provided.

**[0065]** FIG. 5 also illustrates a network 520 that may provide communications for one or more computers accessible by users 510. FIG. 5 also illustrates a data store 550 that

may store data pertaining to various events and metrics for storage devices 540, such as input/output latencies.

**[0066]** A request may be sent to storage analysis function 500 for monitoring and analyzing one or more storage devices. The storage analysis function 500 may be co-hosted with the servers 530 and data store 550, or may reside on devices of users 510, or a combination thereof. In some embodiments, a request may be received from a device of one of the users 510 for storage analysis. In response to receipt of the request, storage analysis function 500 may log the request and provide updates as to the status of the request. The storage analysis function 500 may communicate with other services to facilitate processing of the request, collection of data pertaining to request, and generating interfaces to provide results of the request. The storage analysis function 500 may, for example, provide an interface for facilitating submission of the request. The storage analysis function 500 may further provide an interface for viewing the results of the request.

**[0067]** Storage analysis function 500 may be configured to provide analysis and diagnostics for identifying devices with potential performance issues and/or predicted failure probabilities based on real time or accumulated and/or archived monitoring of various devices. The storage analysis function 500 may access metrics, such as device type, device usage activity, and device latencies. The storage analysis function 500 may be made accessible via an application programming interface (API) or a user interface that may be accessed via a Web browser or other input mechanisms.

**[0068]** In some embodiments, an expert system that utilizes logical inferences based on available information may be used. An expert system may take available information pertaining to actual failures of devices and use the information as input to a rules-based system to generate updated event probabilities. The available information may be provided to, for example, a Bayesian process to determine an updated probability for the event.

**[0069]** Within this operating environment, storage analysis function 500 may determine a predicted probability of a failure or other performance issue. The analysis may be based on the distribution of data, such as a tail latency. Storage analysis function 500 may gather data from other components of the operating environment, such as data store 550. The storage analysis function 500 may also collect information stored in log files and other locations that may be available from computing resources 540. The information may also be obtained by querying devices for data that is not currently being stored in a log file.

**[0070]** After a potential issue has been identified, an alert may be sent to the user or an automated response may be initiated.

**[0071]** In some embodiments, the storage analysis function 500 may monitor failures and device attributes automatically. For example, the storage analysis function 500 may monitor for a number of pre-selected profiles or events, and may do so at a predetermined frequency.

**[0072]** Additionally, storage analysis function 500 may provide users with the ability to request and receive notifications or to take specified actions depending on the failure events. For example, an off-device service in communication with the storage analysis function 500 may send an SMS message to a user to indicate that a device is predicted to fail. A user interface may be provided to allow access to the storage analysis function 500.

**[0073]** In some embodiments, a time frame for the event prediction may be specified as well as the resources to be evaluated. For example, a user may want to determine the predicted failure rate for a group of devices over a period of one month.

**[0074]** The storage analysis function 500 may provide a detailed analysis and summary of results to a requesting user. In some embodiments, an API may be provided to facilitate requests for event information. For example, an API can be called with information such as a device identifier and specified event. After the API is called, in one embodiment the storage analysis function 500 may take actions such as:

- Access previous data for the device.
- Retrieve data of related devices.
- Call available APIs that can provide metrics for the devices.
- Invoke a function to generate a predicted probability of an event.

**[0075]** Using the gathered information, the storage analysis function 500 may analyze the data, combine or aggregate the data or extract portions of the data as appropriate, and invoke a distribution analysis function and/or a probability determination function to generate the predicted probability of an event. The determined event probability may be reported through the API along with details regarding potential causes for the prediction.

**[0076]** In some embodiments, certain parameters of the device can be selected based on biasing to identify parameters that may have greater impact on the failure prediction. For example, a subset of attributes can be weighted to indicate a higher degree of relevance of the attributes to the potential failure of the device.

**[0077]** In some embodiments, an API or a programming interface may be provided for facilitating the submission of storage device analysis requests. The API may be

configured to receive electronic messages that encode identifiers indicative of an analysis request for fulfillment by the storage analysis function 500. In response to receiving one of the electronic messages, the API may send second electronic messages indicative of information indicating that request has been received and may provide the results of the analysis.

**[0078]** Referring to FIG. 6, illustrated is an example implementation of functionality associated with a failure prediction engine 500. The failure prediction engine 500 may access, or cause access to, various data such as stored operation data, metrics, and event histories. The failure prediction engine 500 may also generate or cause the generation of data. Users 610 of the service provider may access a user interface 620 for requesting data. In some embodiments, the user interface 620 can be generated by functions implemented in software executing on one or more servers 635. The requested data may be provided to a data store writer 650 that may store the data. A data store reader 640 may be configured to access the data store 660 and retrieve data based on requests from the users 610 or for other purposes. The failure prediction engine 500 may manage the various data that are to be collected. For example, the failure prediction engine 500 may take actions such as invoking a service to retrieve a set of available metrics pertaining to the request and the associated device. The failure prediction engine 500 may also access activity logs for the device, health status for the device, connection states for the device, configurations for the device, and hardware attributes (e.g., manufacturer, date of manufacture, etc.). Other types of data may be accessed as determined by the failure prediction engine 500, such as the type of device, physical location, installed software, usage patterns, and so on.

**[0079]** Referring to FIG. 7, illustrated is an example system for identifying device performance issues. A user interface (UI) 730 may allow interaction with an application running on computing device 720. The application may interact with an API that enables communication, via a network 710, to storage analysis function 500 that may execute in an environment such as a storage virtualization system 700. The storage analysis function 500 may interact with a plurality of storage devices 740.

Turning now to FIG. 8, illustrated is an example operational procedure for identifying performance issues in a computing environment in accordance with the present disclosure. Such an operational procedure may provide for identifying performance issues and can be provided by functions illustrated, for example, in FIGs. 1-7. The operational procedure may be implemented in a system comprising one or more computing devices. Referring to FIG. 8, operation 800 illustrates collecting performance data for input/output

operations executed at a storage device of a plurality of storage devices of a software-defined storage network. Operation 800 may be followed by operation 801. Operation 801 illustrates receiving, at an interface for managing operational health of the software-defined storage network, a request for a subset of the performance data. In an embodiment, the request may specify:

a plurality of latency ranges, and

a time period for the performance data.

**[0080]** Operation 801 may be followed by operation 803. Operation 803 illustrates responsive to receiving the request, accessing a data store storing thereon the collected performance data. In an embodiment, the collected performance data may include latencies for input/output operations executed at the plurality of storage devices.

**[0081]** Operation 803 may be followed by operation 805. Operation 805 illustrates based on the collected performance data, determining a time-based I/O performance profile for the storage device.

**[0082]** Operation 805 may be followed by operation 807. Operation 807 illustrates determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device.

**[0083]** Operation 807 may be followed by operation 809. Operation 809 illustrates returning, via the interface, data usable to render a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile.

**[0084]** Referring to FIG. 9, illustrated is another example operational procedure for identifying storage device issues in a networked environment. Such an operational procedure may provide for identifying performance issues and can be provided by functions illustrated, for example, in FIGs. 1-8. The operational procedure may be implemented in a system comprising one or more computing devices. Referring to FIG. 9, operation 901 illustrates collecting performance data for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network.

**[0085]** Operation 901 may be followed by operation 903. Operation 903 illustrates based on the collected performance data, determining a time-based I/O performance profile for the storage device.

**[0086]** Operation 903 may be followed by operation 905. Operation 905 illustrates determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device and based on

previously collected performance data for devices of the representative group.

**[0087]** Operation 905 may be followed by operation 907. Operation 907 illustrates determining that a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of a probable failure of the storage device.

**[0088]** Operation 907 may be followed by operation 909. Operation 909 illustrates based on the determining that the storage device exceeded the predetermined deviance threshold, initiating an action to mitigate the probable failure.

**[0089]** The various aspects of the disclosure are described herein with regard to certain examples and embodiments, which are intended to illustrate but not to limit the disclosure. It should be appreciated that the subject matter presented herein may be implemented as a computer process, a computer-controlled apparatus, or a computing system or an article of manufacture, such as a computer-readable storage medium. While the subject matter described herein is presented in the general context of program modules that execute on one or more computing devices, those skilled in the art will recognize that other implementations may be performed in combination with other types of program modules. Generally, program modules include routines, programs, components, data structures and other types of structures that perform particular tasks or implement particular abstract data types.

**[0090]** Those skilled in the art will also appreciate that the subject matter described herein may be practiced on or in conjunction with other computer system configurations beyond those described herein, including multiprocessor systems. The embodiments described herein may also be practiced in distributed computing environments, where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote memory storage devices.

**[0091]** Networks established by or on behalf of a user to provide one or more services (such as various types of cloud-based computing or storage) accessible via the Internet and/or other networks to a distributed set of clients may be referred to as a service provider. Such a network may include one or more data centers such as data center 100 illustrated in FIG. 1, which are configured to host physical and/or virtualized computer servers, storage devices, networking equipment and the like, that may be used to implement and distribute the infrastructure and services offered by the service provider.

**[0092]** In some embodiments, a server that implements a portion or all of one or

more of the technologies described herein, including the techniques to implement the capturing of network traffic may include a general-purpose computer system that includes or is configured to access one or more computer-accessible media. FIG. 10 illustrates such a general-purpose computing device 1000. In the illustrated embodiment, computing device 1000 includes one or more processors 1010a, 1010b, and/or 1010n (which may be referred herein singularly as "a processor 1010" or in the plural as "the processors 1010") coupled to a system memory 1020 via an input/output (I/O) interface 1030. Computing device 1000 further includes a network interface 1040 coupled to I/O interface 1030.

**[0093]** In various embodiments, computing device 1000 may be a uniprocessor system including one processor 1010 or a multiprocessor system including several processors 1010 (e.g., two, four, eight, or another suitable number). Processors 1010 may be any suitable processors capable of executing instructions. For example, in various embodiments, processors 1010 may be general-purpose or embedded processors implementing any of a variety of instruction set architectures (ISAs), such as the x106, PowerPC, SPARC, or MIPS ISAs, or any other suitable ISA. In multiprocessor systems, each of processors 1010 may commonly, but not necessarily, implement the same ISA.

**[0094]** System memory 1020 may be configured to store instructions and data accessible by processor(s) 1010. In various embodiments, system memory 1020 may be implemented using any suitable memory technology, such as static random access memory (SRAM), synchronous dynamic RAM (SDRAM), nonvolatile/Flash-type memory, or any other type of memory. In the illustrated embodiment, program instructions and data implementing one or more desired functions, such as those methods, techniques and data described above, are shown stored within system memory 1020 as code 1025 and data 1026.

**[0095]** In one embodiment, I/O interface 1030 may be configured to coordinate I/O traffic between the processor 1010, system memory 1020, and any peripheral devices in the device, including network interface 1040 or other peripheral interfaces. In some embodiments, I/O interface 1030 may perform any necessary protocol, timing, or other data transformations to convert data signals from one component (e.g., system memory 1020) into a format suitable for use by another component (e.g., processor 1010). In some embodiments, I/O interface 1030 may include support for devices attached through various types of peripheral buses, such as a variant of the Peripheral Component Interconnect (PCI) bus standard or the Universal Serial Bus (USB) standard, for example. In some embodiments, the function of I/O interface 1030 may be split into two or more separate components. Also, in some embodiments some or all of the functionality of I/O interface

1030, such as an interface to system memory 1020, may be incorporated directly into processor 1010.

**[0096]** Network interface 1040 may be configured to allow data to be exchanged between computing device 1000 and other device or devices 1060 attached to a network or network(s)1050, such as other computer systems or devices as illustrated in FIGS. 1 through 4, for example. In various embodiments, network interface 1040 may support communication via any suitable wired or wireless general data networks, such as types of Ethernet networks, for example. Additionally, network interface 1040 may support communication via telecommunications/telephony networks such as analog voice networks or digital fiber communications networks, via storage area networks such as Fibre Channel SANs or via any other suitable type of network and/or protocol.

**[0097]** In some embodiments, system memory 1020 may be one embodiment of a computer-accessible medium configured to store program instructions and data as described above for FIGS. 1-7 for implementing embodiments of the corresponding methods and apparatus. However, in other embodiments, program instructions and/or data may be received, sent or stored upon different types of computer-accessible media. A computer-accessible medium may include non-transitory storage media or memory media, such as magnetic or optical media, e.g., disk or DVD/CD coupled to computing device 1000 via I/O interface 1030. A non-transitory computer-accessible storage medium may also include any volatile or non-volatile media, such as RAM (e.g. SDRAM, DDR SDRAM, RDRAM, SRAM, etc.), ROM, etc., that may be included in some embodiments of computing device 1000 as system memory 1020 or another type of memory. Further, a computer-accessible medium may include transmission media or signals such as electrical, electromagnetic or digital signals, conveyed via a communication medium such as a network and/or a wireless link, such as may be implemented via network interface 1040. Portions or all of multiple computing devices, such as those illustrated in FIG. 10, may be used to implement the described functionality in various embodiments; for example, software components running on a variety of different devices and servers may collaborate to provide the functionality. In some embodiments, portions of the described functionality may be implemented using storage devices, network devices, or special-purpose computer systems, in addition to or instead of being implemented using general-purpose computer systems. The term "computing device," as used herein, refers to at least all these types of devices and is not limited to these types of devices.

**[0098]** Various storage devices and their associated computer-readable media

provide non-volatile storage for the computing devices described herein. Computer-readable media as discussed herein may refer to a mass storage device, such as a solid-state drive, a hard disk or CD-ROM drive. However, it should be appreciated by those skilled in the art that computer-readable media can be any available computer storage media that can be accessed by a computing device.

**[099]** By way of example, and not limitation, computer storage media may include volatile and non-volatile, removable and non-removable media implemented in any method or technology for storage of information such as computer-readable instructions, data structures, program modules or other data. For example, computer media includes, but is not limited to, RAM, ROM, EPROM, EEPROM, flash memory or other solid state memory technology, CD-ROM, digital versatile disks ("DVD"), HD-DVD, BLU-RAY, or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by the computing devices discussed herein. For purposes of the claims, the phrase "computer storage medium," "computer-readable storage medium" and variations thereof, does not include waves, signals, and/or other transitory and/or intangible communication media, per se.

**[0100]** Encoding the software modules presented herein also may transform the physical structure of the computer-readable media presented herein. The specific transformation of physical structure may depend on various factors, in different implementations of this description. Examples of such factors may include, but are not limited to, the technology used to implement the computer-readable media, whether the computer-readable media is characterized as primary or secondary storage, and the like. For example, if the computer-readable media is implemented as semiconductor-based memory, the software disclosed herein may be encoded on the computer-readable media by transforming the physical state of the semiconductor memory. For example, the software may transform the state of transistors, capacitors, or other discrete circuit elements constituting the semiconductor memory. The software also may transform the physical state of such components in order to store data thereupon.

**[0101]** As another example, the computer-readable media disclosed herein may be implemented using magnetic or optical technology. In such implementations, the software presented herein may transform the physical state of magnetic or optical media, when the software is encoded therein. These transformations may include altering the magnetic characteristics of particular locations within given magnetic media. These transformations

also may include altering the physical features or characteristics of particular locations within given optical media, to change the optical characteristics of those locations. Other transformations of physical media are possible without departing from the scope and spirit of the present description, with the foregoing examples provided only to facilitate this discussion.

**[0102]** In light of the above, it should be appreciated that many types of physical transformations take place in the disclosed computing devices in order to store and execute the software components and/or functionality presented herein. It is also contemplated that the disclosed computing devices may not include all of the illustrated components shown in FIG. 10, may include other components that are not explicitly shown in FIG. 10, or may utilize an architecture completely different than that shown in FIG. 10.

**[0103]** Although the various configurations have been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended representations is not necessarily limited to the specific features or acts described. Rather, the specific features and acts are disclosed as example forms of implementing the claimed subject matter.

**[0104]** Conditional language used herein, such as, among others, "can," "could," "might," "may," "e.g.," and the like, unless specifically stated otherwise, or otherwise understood within the context as used, is generally intended to convey that certain embodiments include, while other embodiments do not include, certain features, elements, and/or steps. Thus, such conditional language is not generally intended to imply that features, elements, and/or steps are in any way required for one or more embodiments or that one or more embodiments necessarily include logic for deciding, with or without author input or prompting, whether these features, elements, and/or steps are included or are to be performed in any particular embodiment. The terms "comprising," "including," "having," and the like are synonymous and are used inclusively, in an open-ended fashion, and do not exclude additional elements, features, acts, operations, and so forth. Also, the term "or" is used in its inclusive sense (and not in its exclusive sense) so that when used, for example, to connect a list of elements, the term "or" means one, some, or all of the elements in the list.

**[0105]** While certain example embodiments have been described, these embodiments have been presented by way of example only, and are not intended to limit the scope of the inventions disclosed herein. Thus, nothing in the foregoing description is intended to imply that any particular feature, characteristic, step, module, or block is

necessary or indispensable. Indeed, the novel methods and systems described herein may be embodied in a variety of other forms; furthermore, various omissions, substitutions and changes in the form of the methods and systems described herein may be made without departing from the spirit of the inventions disclosed herein. The accompanying claims and their equivalents are intended to cover such forms or modifications as would fall within the scope and spirit of certain of the inventions disclosed herein.

**[0106]** It should be appreciated any reference to "first," "second," etc. items and/or abstract concepts within the description is not intended to and should not be construed to necessarily correspond to any reference of "first," "second," etc. elements of the claims. In particular, within this Summary and/or the following Detailed Description, items and/or abstract concepts such as, for example, individual computing devices and/or operational states of the computing cluster may be distinguished by numerical designations without such designations corresponding to the claims or even other paragraphs of the Summary and/or Detailed Description. For example, any designation of a "first operational state" and "second operational state" of the computing cluster within a paragraph of this disclosure is used solely to distinguish two different operational states of the computing cluster within that specific paragraph – not any other paragraph and particularly not the claims.

**[0107]** In closing, although the various techniques have been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended representations is not necessarily limited to the specific features or acts described. Rather, the specific features and acts are disclosed as example forms of implementing the claimed subject matter.

**[0108]** The disclosure presented herein encompasses the subject matter set forth in the following example clauses.

**[0109]** A computer-implemented method comprising:

**[0110]** a. receiving a request for input/output performance data for a plurality of storage devices of a software-defined storage network, the request specifying:

**[0111]** 1. a plurality of latency ranges, and

**[0112]** 2. a time period for the input/output performance data;

**[0113]** In an embodiment, a client can request data for the total number of I/O operations where the latency for the operations exceeded specified latency thresholds such as greater than 256us, 1ms, 4ms, etc.; the client can request a time period for the I/O latency data, such as over a two-week period.

**[0114]** b. responsive to receiving the request at the interface, accessing a data store

storing thereon a historical record of latencies for input/output operations executed at the plurality of storage devices; in an embodiment, a database stores the actual latencies that each I/O operation that was executed at each storage device of the storage network;

**[0115]** c. determining a number of latencies that fall into each of the specified latency ranges and within the specified time period; in an embodiment, a function analyzes the latency records and identifies latencies that exceeded each of the specified latency thresholds and fell within the specified time period;

**[0116]** d. returning, via the interface, data usable to render a histogram indicative of the number of latencies that fall into each of the specified latency ranges and within the specified time period; in an embodiment, the user interface returns data for generating a histogram for the number of latencies that exceeded each of the specified latency thresholds and that occurred during the specified time period.

**[0117]** A computer-implemented method comprising:

**[0118]** a. receiving, at an interface for managing operational health of a plurality of storage devices of a software-defined storage network, a request including an identifier of a selected storage device of the software-defined storage network; in an embodiment, a user interface is instantiated that operates to receive, from a requesting process, an identifier of a selected storage device of the software-defined storage network;

**[0119]** b. responsive to receiving the request at the interface, accessing a data store storing thereon a historical record of latencies for input/output operations executed at the selected storage device of the software-defined storage network; in an embodiment, a database stores the actual latencies for each I/O operation that was executed at each storage device of the storage network;

**[0120]** c. determining a latency profile for the selected storage device, the latency profile determined based on a number of latencies for the selected storage device that fall within each of a plurality of latency ranges and within a selected time period; in an embodiment, a latency profile includes the distribution of the numbers of latencies that fell within each of a plurality of latency ranges (e.g., the total number of I/O operations where the latency was greater than 256us, 1ms, 4ms, etc.) and within a selected time period data (e.g., during the previous two weeks).]

**[0121]** d. determining that a difference between the latency profile for the selected storage device and a characteristic latency profile for a class of storage devices of the software-defined storage network exceeds a predetermined deviance threshold; in an embodiment, a function determines that the difference between the latency profile and a

characteristic latency profile for a class of storage devices (e.g., a group of storage devices with the same type of storage such as DIMM and same capacity) exceeds a predetermined deviance threshold (e.g., the highest latency of the device occurred more than 1% of the time and was greater than two standard deviations from the characteristic profile for the class of devices).

**[0122]** e. returning, via the interface, data usable to render an indication that the selected storage device exceeded the predetermined deviance threshold; in an embodiment, the interface returns a notification that the selected storage device exhibited a latency profile that deviated from the characteristic profile of similar devices.

**[0123]** The disclosure presented herein encompasses the subject matter set forth in the following example clauses.

**[0124]** A computer-implemented method comprising:

receiving a request for input/output performance data for a plurality of storage devices of a software-defined storage network, the request specifying:

a plurality of latency ranges, and

a time period for the input/output performance data;

responsive to receiving the request at the interface, accessing a data store storing thereon a historical record of latencies for input/output operations executed at the plurality of storage devices;

determining a number of input/output operations with latencies that fall into each of the specified latency ranges and within the specified time period;

returning, via the interface, data usable to render:

a histogram indicative of the number of latencies that fall into each of the specified latency ranges and within the specified time period; and

an indication that one or more of the specified latency ranges have exceeded a threshold indicated of a failure.

**[0125]** A computer-implemented method comprising:

receiving, at an interface for managing operational health of a plurality of storage devices of a software-defined storage network, a request including an identifier of a selected storage device of the software-defined storage network;

responsive to receiving the request at the interface, accessing a data store storing thereon a historical record of latencies for input/output operations executed at the selected storage device of the software-defined storage network;

determining a latency profile for the selected storage device, the latency profile

determined based on a number of latencies for the selected storage device that fall within each of a plurality of latency ranges and within a selected time period;

determining that a difference between the latency profile for the selected storage device and a characteristic latency profile for a class of storage devices of the software-defined storage network exceeds a predetermined deviance threshold; and

returning, via the interface, data usable to render an indication that the selected storage device exceeded the predetermined deviance threshold.

EXAMPLE CLAUSES

**[0126]** The disclosure presented herein encompasses the subject matter set forth in the following example clauses.

**[0127]** Example Clause A, a computer-implemented method for operating a storage network, the method comprising:

instantiating an interface for managing operational health of a software-defined storage network, wherein the interface is configured to enable selection of a range of input/output latencies ranges and a time period for performance data for input/output operations executed at a selected storage device of a plurality of storage devices of the storage network;

receiving a request for performance data collected for input/output operations executed at a selected storage device of the plurality of storage devices, the request including a selected latency range and a selected time period;

responsive to receiving the request, accessing a data store storing thereon performance data collected for input/output operations executed at the plurality of storage devices;

based on the accessed performance data and the selected latency range and selected time period, determining a time-based I/O performance profile for the selected storage device and a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the selected storage device; and

returning, via the interface:

data usable to enable a comparison between the time-based I/O performance profile for the selected storage device and the characteristic time-based I/O performance profile; and

data usable to render an indication of whether the time-based I/O performance profile for the selected storage device is indicative of a probable fault.

**[0128]** Example Clause B, the computer-implemented method of Example Clause A,

further comprising returning, via the interface, data usable to render a deviance threshold that is determined based on a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile.

**[0129]** Example Clause C, the computer-implemented method of any one of Example Clauses A through B, wherein the performance data comprises latencies for I/O operations that were executed at the storage device.

**[0130]** Example Clause D, the computer-implemented method of any one of Example Clauses A through C, wherein the time-based I/O performance profile comprises a distribution of numbers of metrics that fell within each of a plurality of metric ranges.

**[0131]** Example Clause E, the computer-implemented method of any one of Example Clauses A through D, wherein the time-based I/O performance profile comprises.

**[0132]** Example Clause F, the computer-implemented method of any one of Example Clauses A through E, wherein the data usable to render a histogram indicative of a count of latencies that fall into latencies of the selected latency range and within the selected time period.

**[0133]** While Example Clauses A through F are described above with respect to a computer-implemented method, it is understood in the context of this disclosure that the subject matter of Example Clauses A through F can additionally or alternatively be implemented by a system or device or computer readable medium.

**[0134]** Example Clause G, a computing device comprising:

one or more processors;

a memory in communication with the one or more processors, the memory having computer-readable instructions stored thereupon which, when executed by the one or more processors, cause the computing device perform operations comprising:

receiving a request for performance data for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network, the request including a selected latency range and a selected time period;

responsive to receiving the request, accessing a data store storing thereon a historical record of latencies for input/output operations executed at the plurality of storage devices;

based on the accessed data, determining a time-based I/O performance profile for the storage device that is based on one or more performance metrics for the storage device at the selected latency range and selected time period;

determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device;

returning data usable to render:

a visual representation of the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile; and

a visual representation of a predetermined deviance threshold that is indicative of the storage device entering a marginal performance state.

**[0135]** Example Clause H, the system of Example Clause G, wherein data further comprises data usable to render a probability that the storage device will enter a marginal performance state exceeds a predetermined threshold.

**[0136]** Example Clause I, the system of any one of Example Clauses G through H, wherein the probability is determined in part on a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile for the representative group of storage devices.

**[0137]** Example Clause J, the system of any one of Example Clauses G through I, wherein the request includes a plurality of latency ranges, and a time period for the input/output operations.

**[0138]** Example Clause K, the system of any one of Example Clauses G through J, wherein the time-based I/O performance profile comprises a total number of I/O operations where associated latencies were greater than a threshold.

**[0139]** Example Clause L, the system of any one of Example Clauses G through K, wherein the plurality of storage devices is an erasure coding group or a group of mirrored devices.

**[0140]** Example Clause M, the system of any one of Example Clauses G through L, wherein the data usable to render a histogram indicative of a number of latencies that fall into each of specified latency ranges and within a specified time period.

**[0141]** Example Clause N, the system of any one of Example Clauses G through M, wherein the probability is determined in part on a time-based distribution of latencies for I/O operations associated with the storage device.

**[0142]** Example Clause O, the system of any one of Example Clauses G through N, wherein the probability is determined in part on latencies for I/O operations associated with the storage device that have exceeded a threshold value.

**[0143]** Example Clause P, the system of any one of Example Clauses G through O, wherein the time-based I/O performance profile comprises a distribution of numbers of metrics that fell within each of a plurality of metric ranges.

**[0144]** Example Clause Q, the system of any one of Example Clauses G through P,

wherein the determining that the storage device is no longer in the marginal performance state is based on determining that latencies for I/O operations associated with the storage device have not exceeded a predetermined threshold for a specified time period.

**[0145]** While Example Clauses G through Q are described above with respect to a system, it is understood in the context of this disclosure that the subject matter of Example Clauses G through Q can additionally or alternatively be implemented by a device3 or method or computer readable medium.

**[0146]** Example Clause R, a computer-readable medium having encoded thereon computer-executable instructions that, when executed, cause one or more processing units of a computing device to execute a method comprising:

receiving a request for performance data for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network, the request specifying a selected latency range and selected time period;

responsive to receiving the request, accessing a data store storing thereon a historical record of latencies for input/output operations executed at the plurality of storage devices;

based on the accessed data, determining a time-based I/O performance profile for the storage device, the I/O performance profile determined based on performance metrics for the storage device at the selected latency range and selected time period;

determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device; and

returning data usable to render the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile.

**[0147]** Example Clause S, the computer-readable medium of Example Clause R, further comprising returning, via the interface, data usable to render predetermined deviance threshold that is indicative of the storage device entering a marginal performance state.

**[0148]** Example Clause T, the computer-readable medium of any one of Example Clauses R through S, wherein the time-based I/O performance profile comprises a distribution of numbers of metrics that fell within each of a plurality of metric ranges.

**[0149]** While Example Clauses R through T are described above with respect to a computer-readable medium, it is understood in the context of this disclosure that the subject matter of Example Clauses R through T can additionally or alternatively be implemented by a method or via a device or via a system.

**[0150]** The disclosure presented herein encompasses the subject matter set forth in the following example clauses.

**[0151]** Example Clause AA, a computer-implemented method for predicting a state of a storage device, the method comprising:

collecting performance data for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network;

based on the collected performance data, determining a time-based I/O performance profile for the storage device;

determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device and based on previously collected performance data for devices of the representative group;

determining that a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of a probable failure of the storage device; and

based on the determining that the storage device exceeded the predetermined deviance threshold, initiating an action to mitigate the probable failure.

**[0152]** Example Clause BB, the system of Example Clause AA, wherein the plurality of storage devices are configured to implement a fault resiliency function.

**[0153]** Example Clause CC, the system of any one of Example Clauses AA through BB, wherein the plurality of storage devices is an erasure coding group or a group of mirrored devices.

**[0154]** Example Clause DD, the system of any one of Example Clauses AA through CC, wherein the collected performance data comprises latencies for I/O operations that were executed at the storage device.

**[0155]** Example Clause EE, the system of any one of Example Clauses AA through DD, wherein the time-based I/O performance profile comprises a distribution of numbers of metrics that fell within each of a plurality of metric ranges.

**[0156]** Example Clause FF, the system of any one of Example Clauses AA through EE, wherein the time-based I/O performance profile comprises a total number of I/O operations where associated latencies were greater than a threshold.

**[0157]** Example Clause GG, the system of any one of Example Clauses AA through FF, wherein the time-based I/O performance profile for the storage device comprises a latency profile that is determined based on a number of latencies for the storage device that fall within each of a plurality of latency ranges and within a selected time period.

**[0158]** While Example Clauses AA through GG are described above with respect to a

system, it is understood in the context of this disclosure that the subject matter of Example Clauses AA through GG can additionally or alternatively be implemented by a method or device.

**[0159]** Example Clause HH, a computing device comprising:

one or more processors;

a memory in communication with the one or more processors, the memory having computer-readable instructions stored thereupon which, when executed by the one or more processors, cause the computing device perform operations comprising:

determining an I/O performance profile for a storage device, the I/O performance profile determined based on one or more performance metrics for the storage device;

based on the I/O performance profile and a characteristic I/O performance profile for a representative group of storage devices having common characteristics with the storage device, determining that a probability that the storage device will enter a fault state exceeds a predetermined threshold; and

based on the probability that the storage device exceeding the predetermined threshold, identifying the storage device for failure mitigation.

**[0160]** Example Clause II, the system of Example Clause HH, wherein the probability is determined in part on a difference between the I/O performance profile for the storage device and the characteristic I/O performance profile for the representative group of storage devices.

**[0161]** Example Clause JJ, the system of any one of Example Clauses HH through II, wherein the storage device is part of a group of storage devices that are configured to implement a fault resiliency function.

**[0162]** Example Clause KK, the system of any one of Example Clauses HH through JJ, wherein the fault resiliency function is mirroring or erasure coding.

**[0163]** Example Clause LL, the system of any one of Example Clauses HH through KK, wherein the probability is determined in part on a time-based distribution of latencies for I/O operations associated with the storage device.

**[0164]** Example Clause MM, the system of any one of Example Clauses HH through LL, wherein the probability is determined in part on latencies for I/O operations associated with the storage device that have exceeded a threshold value.

**[0165]** Example Clause NN, the system of any one of Example Clauses HH through MM, further comprising determining that the probability that the storage device will enter the fault state has not exceeded the predetermined threshold for a specified time period.

**[0166]** While Example Clauses HH through NN are described above with respect to a system, it is understood in the context of this disclosure that the subject matter of Example Clauses HH through NN can additionally or alternatively be implemented by a method or device or computer readable medium.

**[0167]** Example Clause OO, a computer-readable medium having encoded thereon computer-executable instructions that, when executed, cause one or more processing units of a computing device to execute a method comprising:

collecting performance data for input/output operations executed at a storage device of a group of storage devices of a storage network;

based on the collected performance data, determining a time-based I/O performance profile for the storage device;

determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device;

determining that a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of the storage device entering a fault state; and

in response to determining that the difference exceeds the predetermined deviance threshold, initiating performance of at least one maintenance action for the storage device.

**[0168]** Example Clause PP, the computer-readable medium of Example Clause OO, wherein the performance data comprise latencies for I/O operations that were executed at the group of storage devices.

**[0169]** Example Clause QQ, the computer-readable medium of any one of Example Clauses OO through PP, wherein the time-based I/O performance profile comprises a distribution of numbers of metrics that fell within each of a plurality of metric ranges.

**[0170]** Example Clause RR, the computer-readable medium of any one of Example Clauses OO through QQ, wherein the time-based I/O performance profile comprises a total number of I/O operations where associated latencies were greater than a threshold.

**[0171]** Example Clause SS, the computer-readable medium of any one of Example Clauses OO through RR, wherein the time-based I/O performance profile for the storage device comprises a latency profile is determined based on a number of latencies for the storage device that fall within each of a plurality of latency ranges and within a selected time period.

**[0172]** Example Clause TT, the computer-readable medium of any one of Example

Clauses OO through SS, wherein the group of storage devices are configured to implement a fault resiliency function.

**[0173]** While Example Clauses OO through TT are described above with respect to a computer-readable medium, it is understood in the context of this disclosure that the subject matter of Example Clauses OO through TT can additionally or alternatively be implemented by a method or via a device or via a system.

## CLAIMS

1. A computer-implemented method for predicting a state of a storage device, the method comprising:

collecting performance data for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network;

based on the collected performance data, determining a time-based I/O performance profile for the storage device;

determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device and based on previously collected performance data for devices of the representative group;

determining that a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of a probable failure of the storage device; and

based on the determining that the storage device exceeded the predetermined deviance threshold, initiating an action to mitigate the probable failure.

2. The computer-implemented method of claim 1, wherein the plurality of storage devices are configured to implement a fault resiliency function.

3. The computer-implemented method of claim 1, wherein the plurality of storage devices is an erasure coding group or a group of mirrored devices.

4. The computer-implemented method of claim 1, wherein the collected performance data comprises latencies for I/O operations that were executed at the storage device.

5. The computer-implemented method of claim 1, wherein the time-based I/O performance profile comprises a distribution of numbers of metrics that fell within each of a plurality of metric ranges.

6. The computer-implemented method of claim 1, wherein the time-based I/O performance profile comprises a total number of I/O operations where associated latencies were greater than a threshold.

7. The computer-implemented method of claim 1, wherein the time-based I/O performance profile for the storage device comprises a latency profile that is determined based on a number of latencies for the storage device that fall within each of a plurality of latency ranges and within a selected time period.

8. A computing device comprising:

one or more processors;

a memory in communication with the one or more processors, the memory having computer-readable instructions stored thereupon which, when executed by the one or more processors, cause the computing device perform operations comprising:

determining an I/O performance profile for a storage device, the I/O performance profile determined based on one or more performance metrics for the storage device;

based on the I/O performance profile and a characteristic I/O performance profile for a representative group of storage devices having common characteristics with the storage device, determining that a probability that the storage device will enter a fault state exceeds a predetermined threshold; and

based on the probability that the storage device exceeding the predetermined threshold, identifying the storage device for failure mitigation.

9. The computing device of claim 8, wherein the probability is determined in part on a difference between the I/O performance profile for the storage device and the characteristic I/O performance profile for the representative group of storage devices.

10. The computing device of claim 8, wherein the storage device is part of a group of storage devices that are configured to implement a fault resiliency function.

11. The computing device of claim 10, wherein the fault resiliency function is mirroring or erasure coding.

12. The computing device of claim 8, wherein the probability is determined in part on a time-based distribution of latencies for I/O operations associated with the storage device.

13. The computing device of claim 8, wherein the probability is determined in part on latencies for I/O operations associated with the storage device that have exceeded a threshold value.

14. The computing device of claim 8, further comprising determining that the probability that the storage device will enter the fault state has not exceeded the predetermined threshold for a specified time period.

15. A computer-readable medium having encoded thereon computer-executable instructions that, when executed, cause one or more processing units of a computing device to execute a method comprising:

collecting performance data for input/output operations executed at a storage device of a group of storage devices of a storage network;

based on the collected performance data, determining a time-based I/O performance profile for the storage device;

determining a characteristic time-based I/O performance profile for a representative

group of storage devices having common characteristics with the storage device;

determining that a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of the storage device entering a fault state; and

in response to determining that the difference exceeds the predetermined deviance threshold, initiating performance of at least one maintenance action for the storage device.

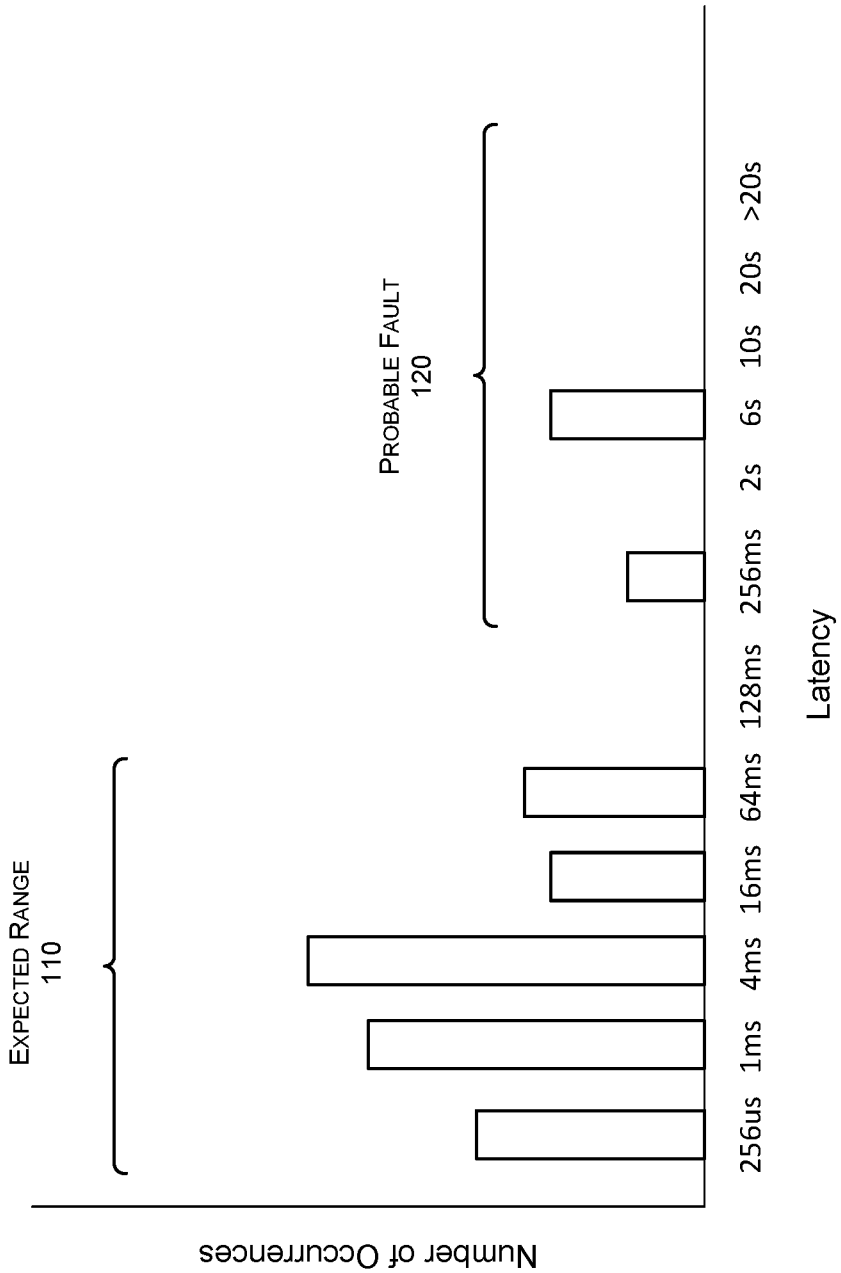

Number of Occurrences
Latency
256us
1ms
4ms
16ms
64ms
128ms
256ms
2s
6s
10s
20s
>20s
Expected Range
110
Probable Fault
120

FIG. 1

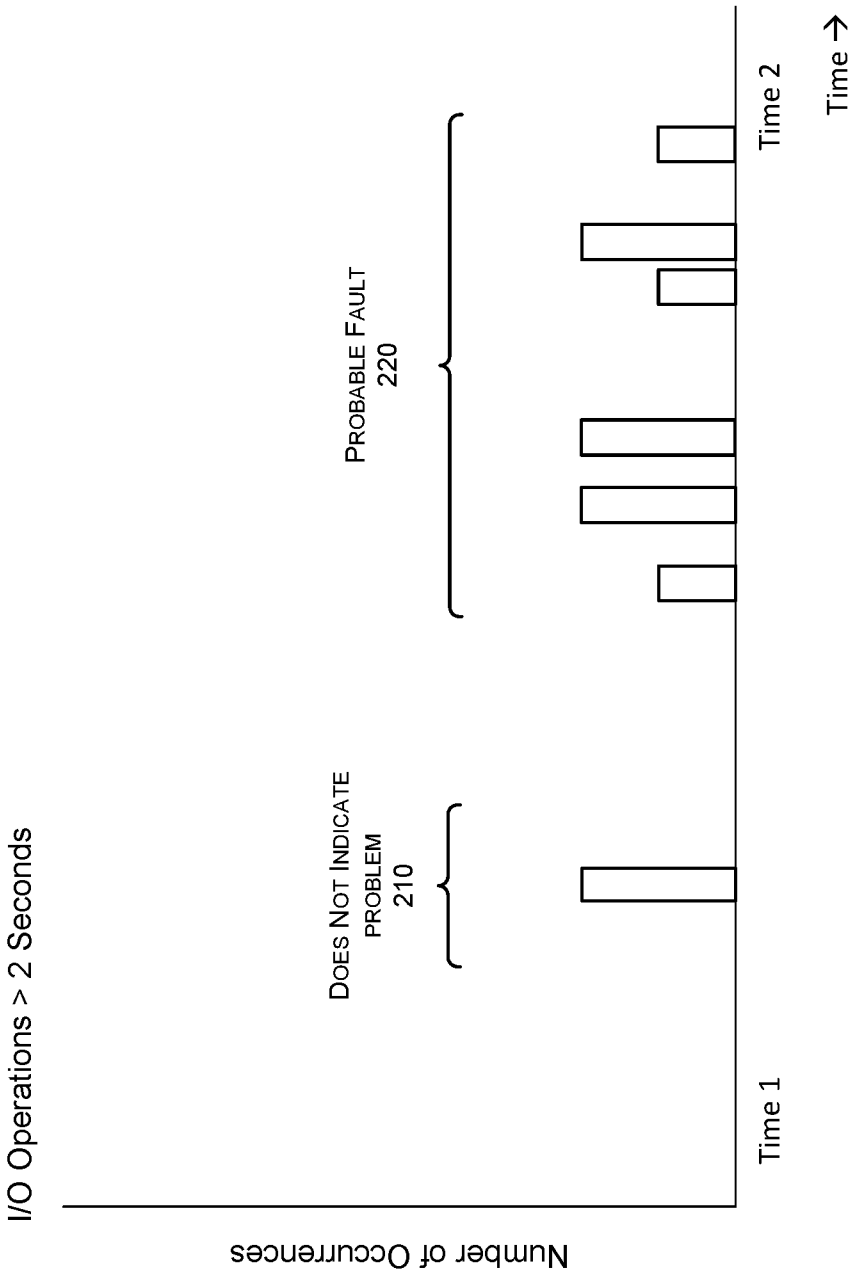

I/O Operations > 2 Seconds
Number of Occurrences
Time 1
Time 2
Time →
Does Not Indicate Problem 210
Probable Fault 220

FIG. 2A

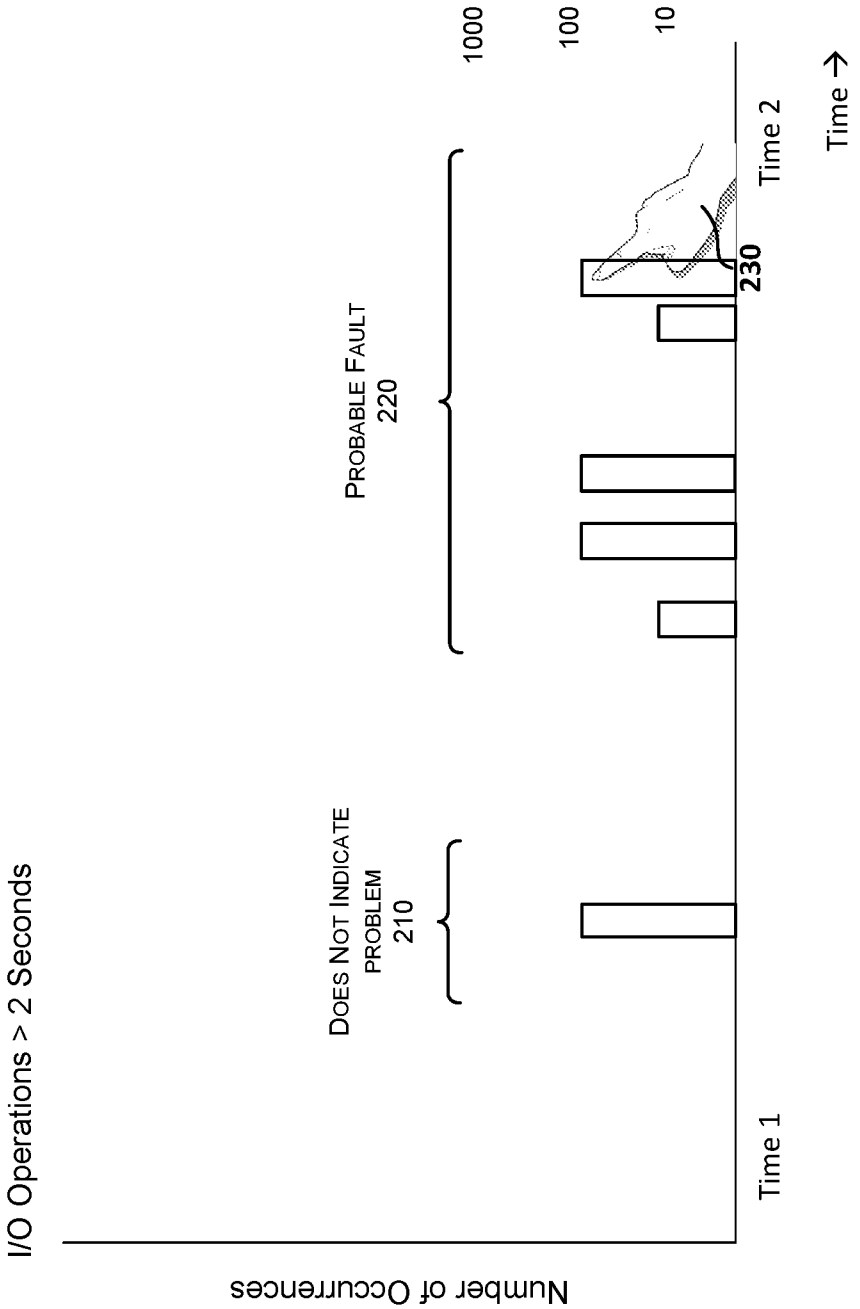

I/O Operations > 2 Seconds
Number of Occurrences
1000
100
10
Time 1
Time 2
Time →
Does Not Indicate Problem 210
Probable Fault 220
230

FIG. 2B

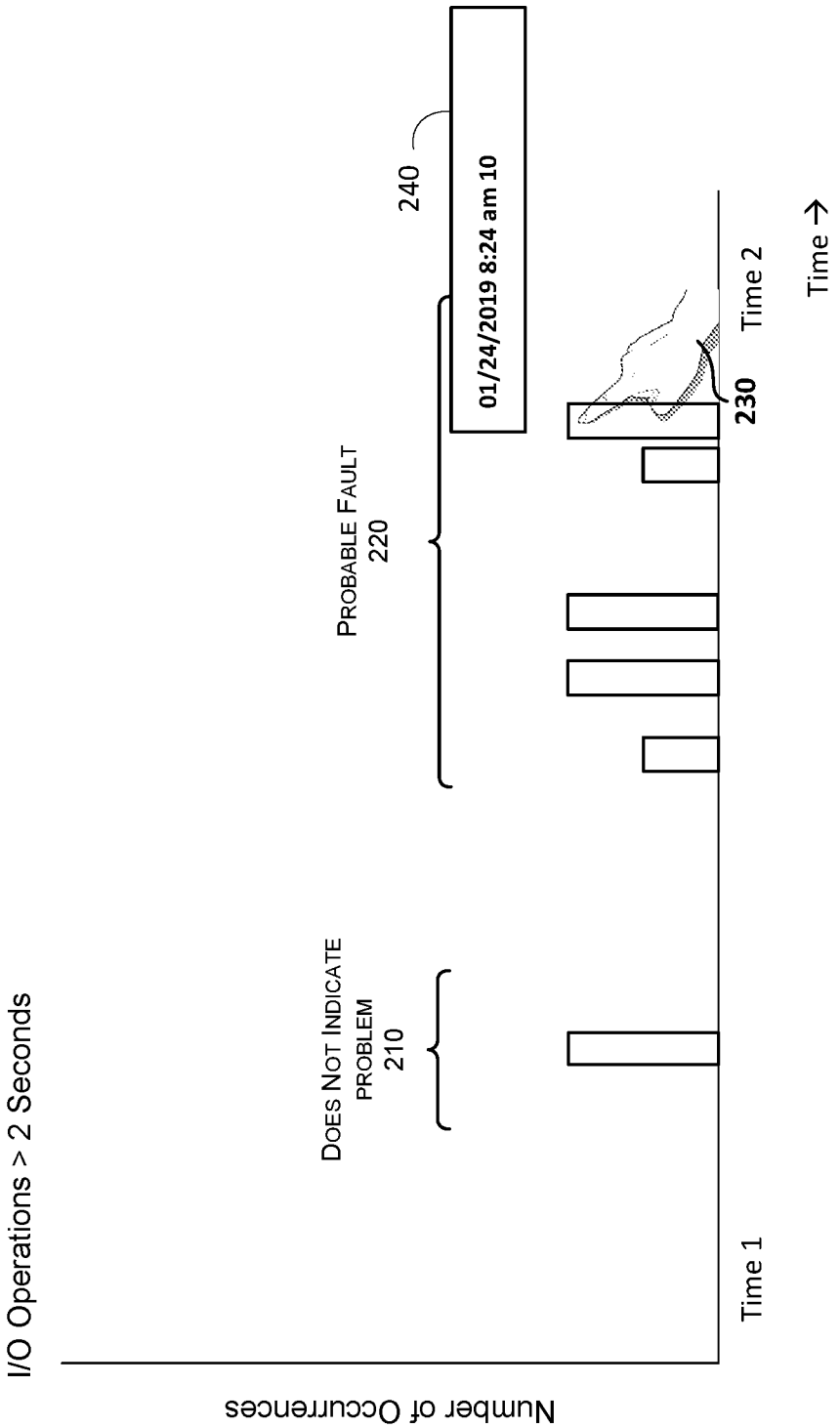

I/O Operations > 2 Seconds
Number of Occurrences
Time 1
Time 2
Time →
Does Not Indicate Problem 210
Probable Fault 220
01/24/2019 8:24 am 10
240
230

FIG. 2C

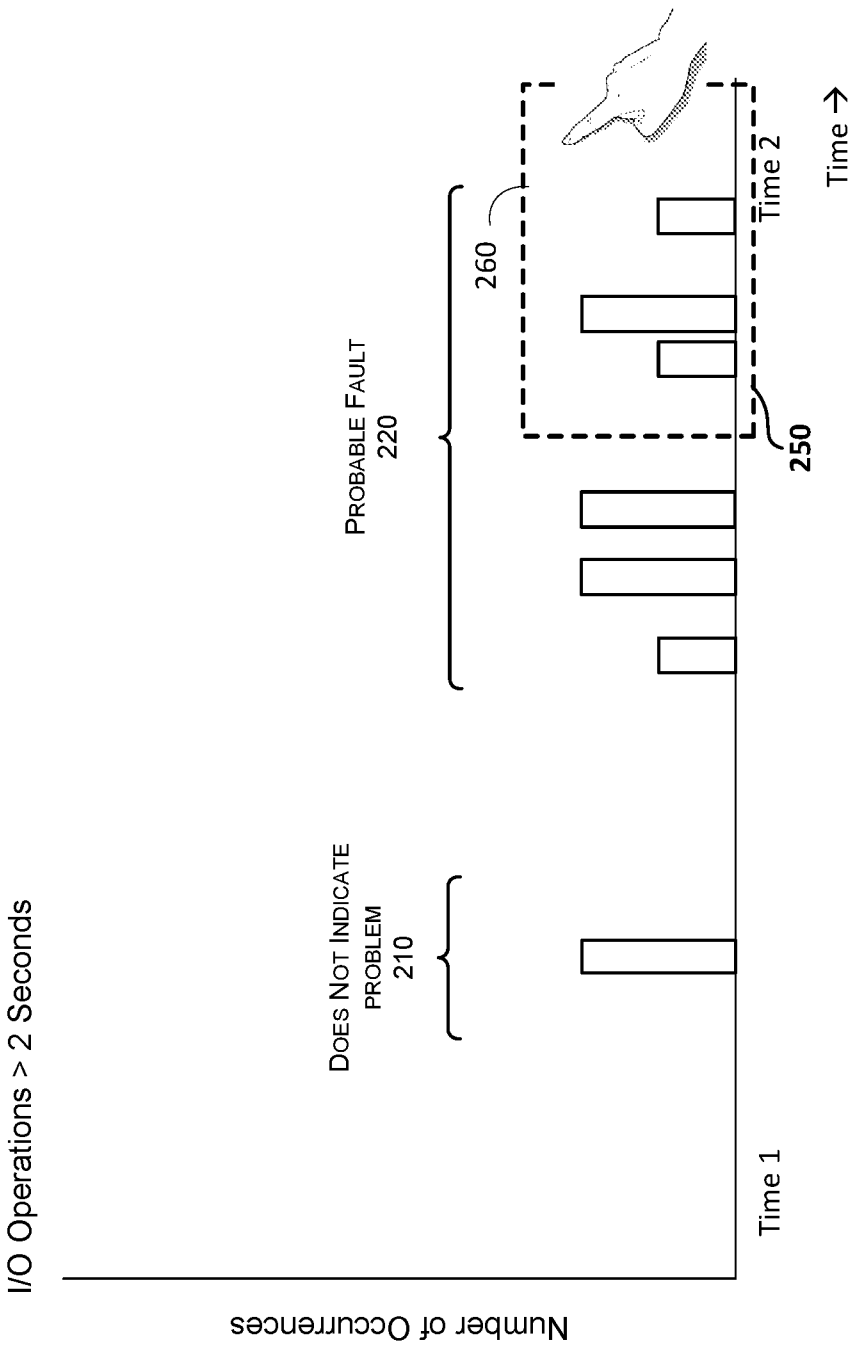

I/O Operations > 2 Seconds
Number of Occurrences
Time 1
Time 2
Time →
Does Not Indicate Problem 210
Probable Fault 220
260
250

FIG. 2D

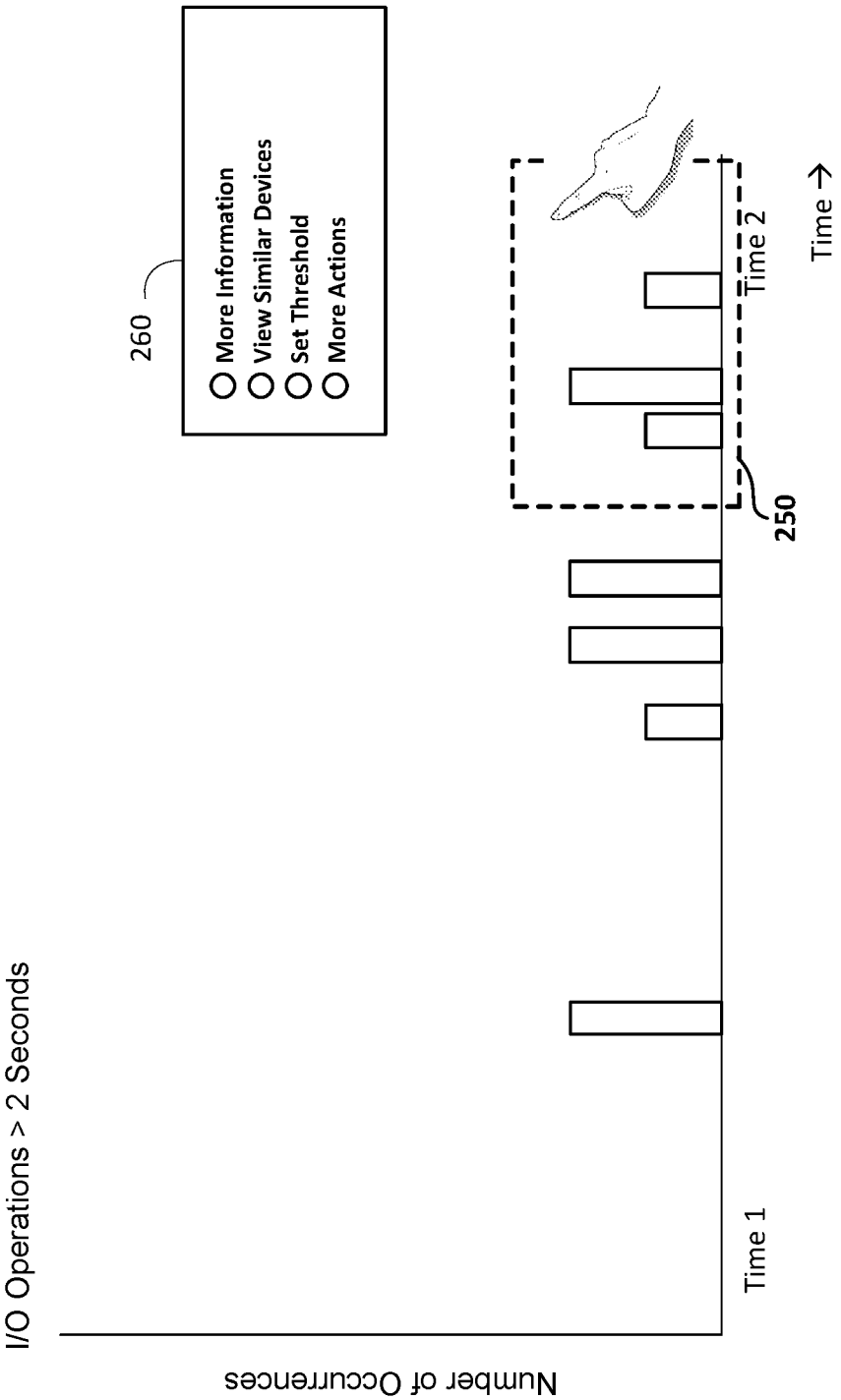

260
More Information
View Similar Devices
Set Threshold
More Actions
Time →
Time 2
250
Time 1
I/O Operations > 2 Seconds
Number of Occurrences

FIG. 2E

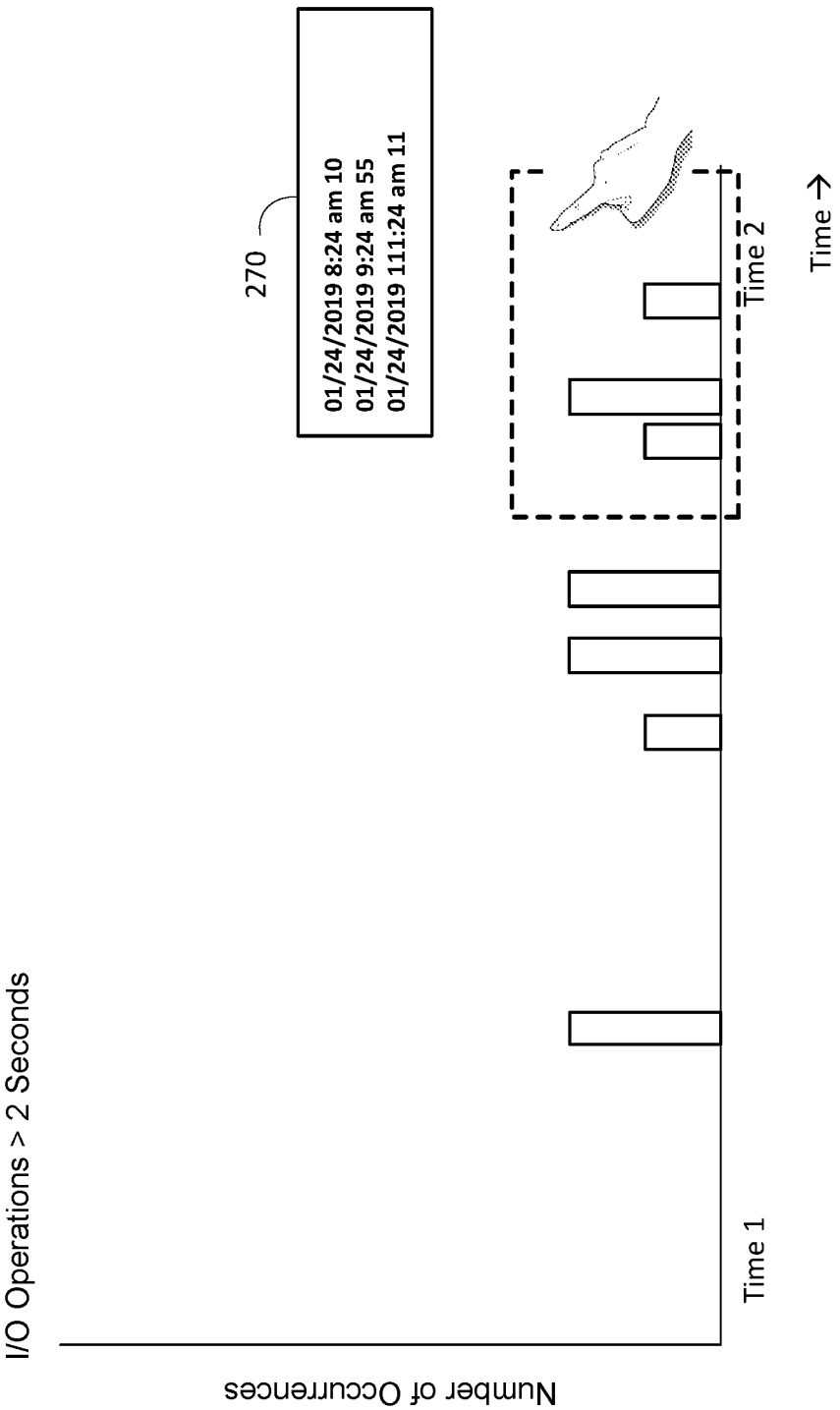

270
01/24/2019 8:24 am 10
01/24/2019 9:24 am 55
01/24/2019 111:24 am 11
Time 2
Time →
Time 1
I/O Operations > 2 Seconds
Number of Occurrences

FIG. 2F

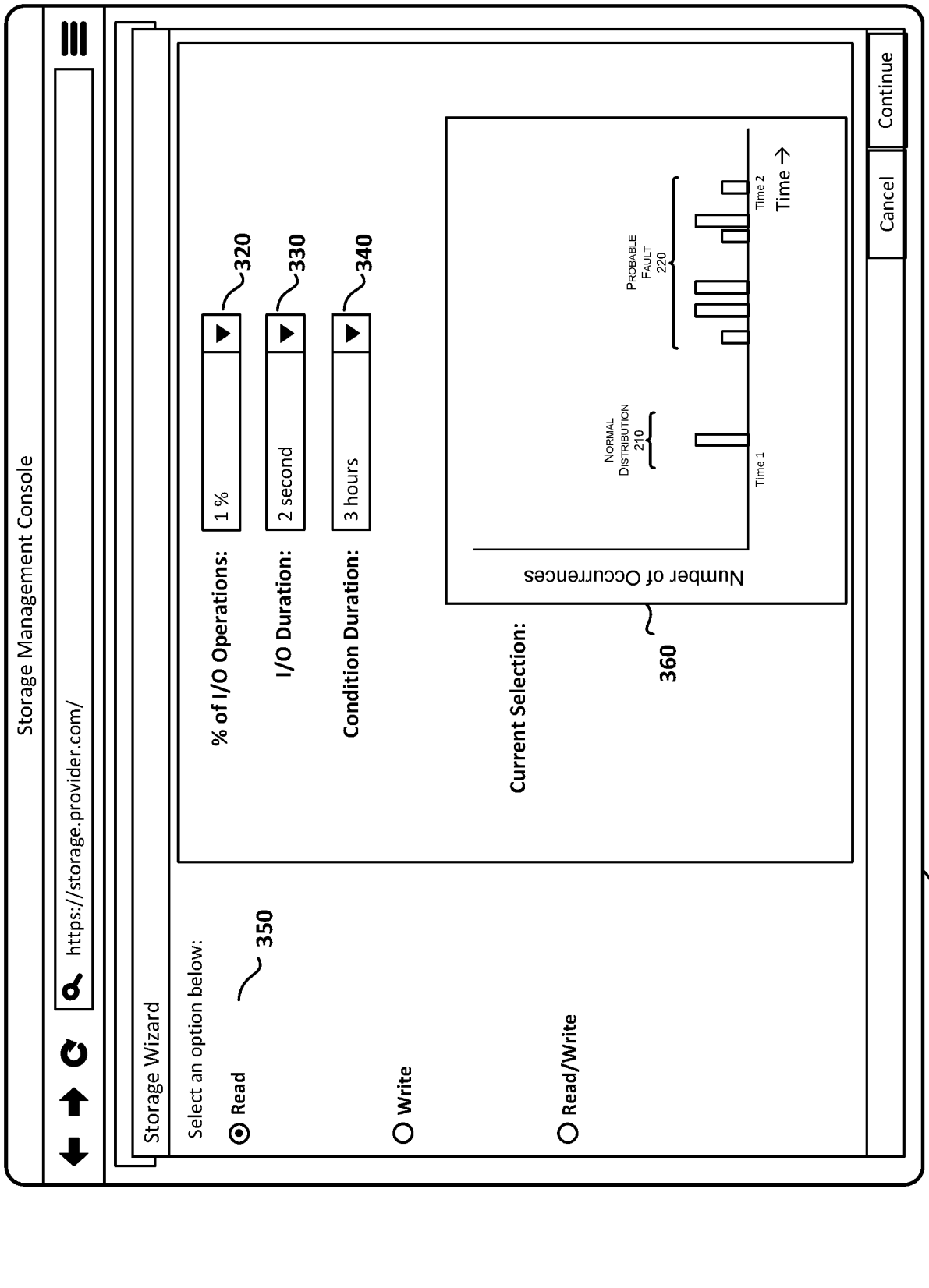

Storage Management Console
https://storage.provider.com/
Storage Wizard
Select an option below:
Read
Write
Read/Write
350
% of I/O Operations:
1 %
320
I/O Duration:
2 second
330
Condition Duration:
3 hours
340
Current Selection:
Number of Occurrences
Normal Distribution 210
Probable Fault 220
Time 1
Time 2
Time →
360
Cancel
Continue
310
300

FIG. 3A

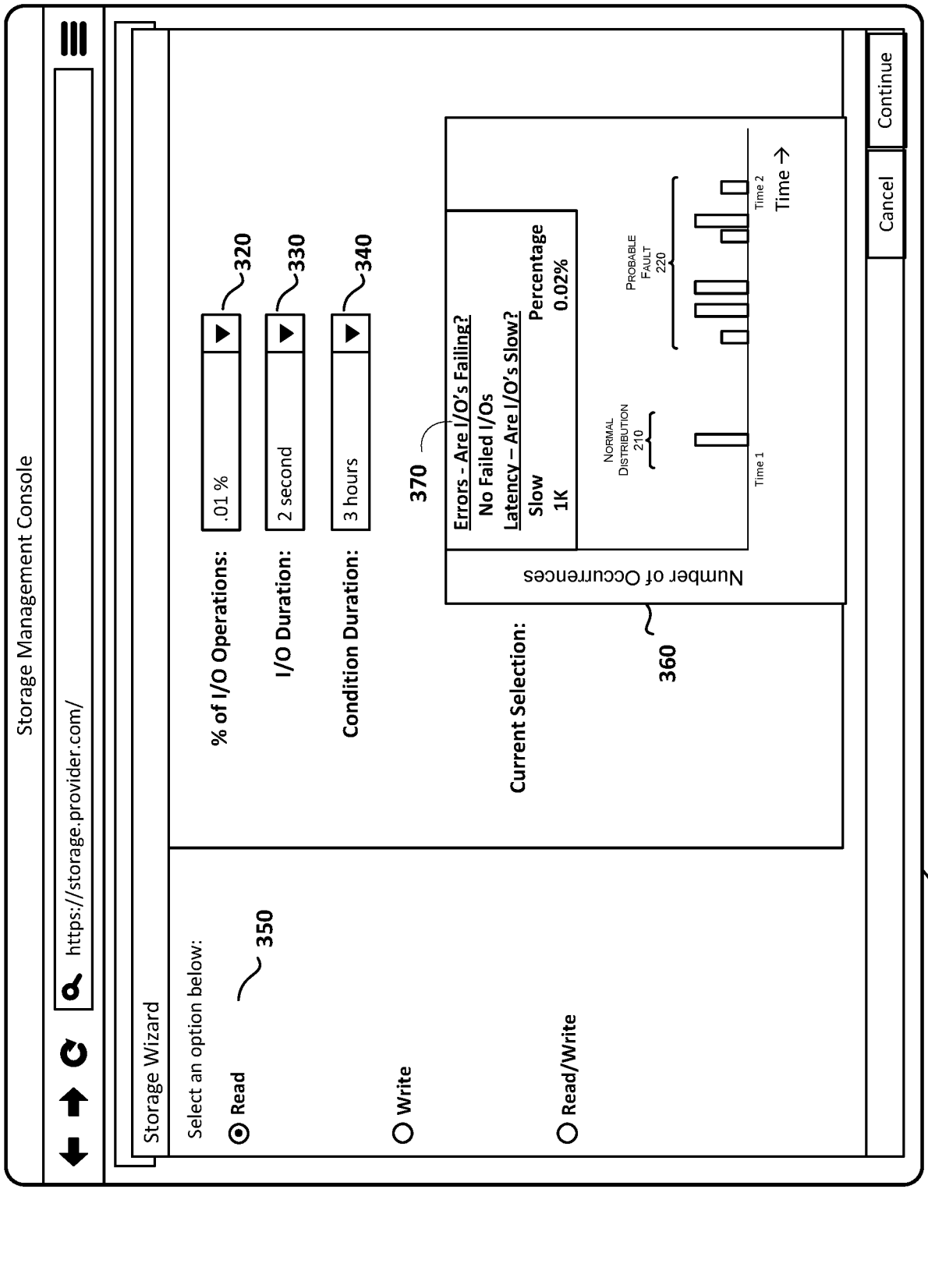

Storage Management Console
https://storage.provider.com/
Storage Wizard
Select an option below:
Read
Write
Read/Write
350
% of I/O Operations:
I/O Duration:
Condition Duration:
.01 %
2 second
3 hours
320
330
340
370
Errors - Are I/O's Failing?
No Failed I/Os
Latency – Are I/O's Slow?
Slow
Percentage
1K
0.02%
Current Selection:
360
Number of Occurrences
Normal Distribution 210
Probable Fault 220
Time 1
Time 2
Time →
Cancel
Continue
310
300

FIG. 3B

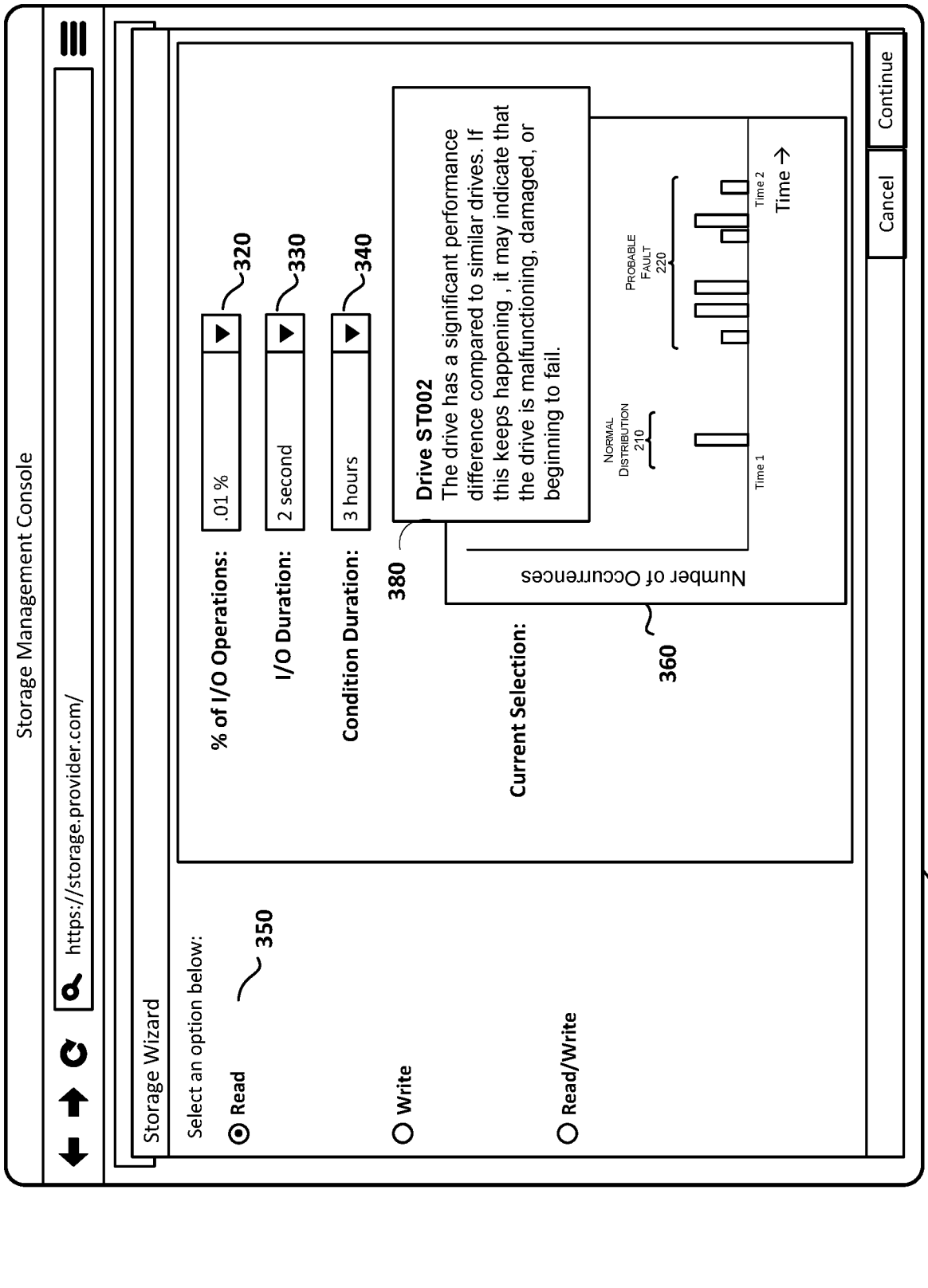

Storage Management Console
https://storage.provider.com/
Storage Wizard
Select an option below:
Read
Write
Read/Write
350
% of I/O Operations:
.01 %
320
I/O Duration:
2 second
330
Condition Duration:
3 hours
340
380
Drive ST002
The drive has a significant performance difference compared to similar drives. If this keeps happening , it may indicate that the drive is malfunctioning, damaged, or beginning to fail.
Current Selection:
360
Number of Occurrences
Normal Distribution 210
Probable Fault 220
Time 1
Time 2
Time →
Cancel
Continue
310
300

FIG. 3C

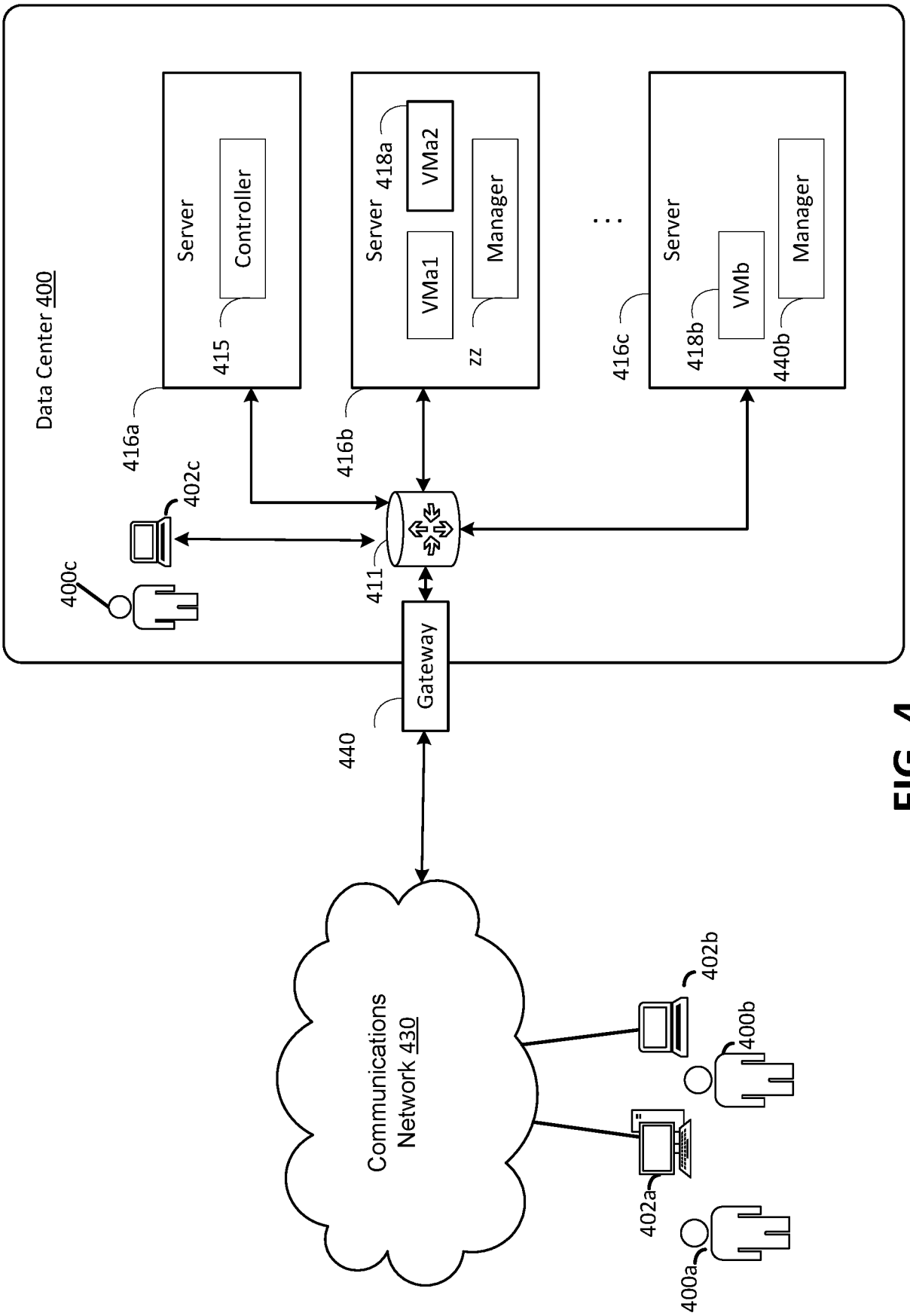

Data Center 400
Server
Controller
415
416a
Server 418a
VMa2
VMa1
Manager
zz
416b
Server
VMb
Manager
418b
440b
416c
402c
400c
411
Gateway
440
Communications Network 430
402b
400b
402a
400a

FIG. 4

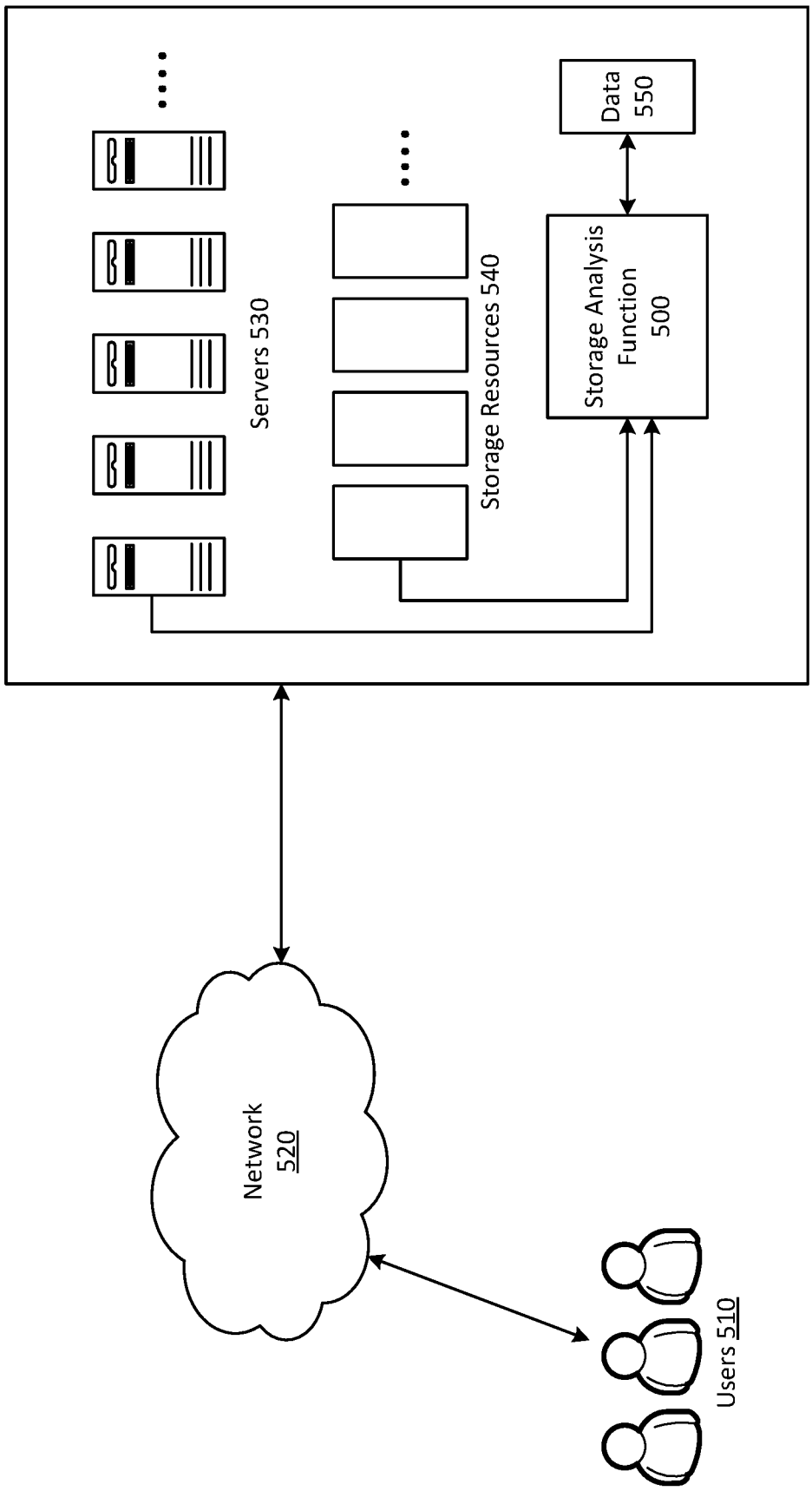

Data
550
Storage Analysis
Function
500
Servers 530
Storage Resources 540
Network
520
Users 510

FIG. 5

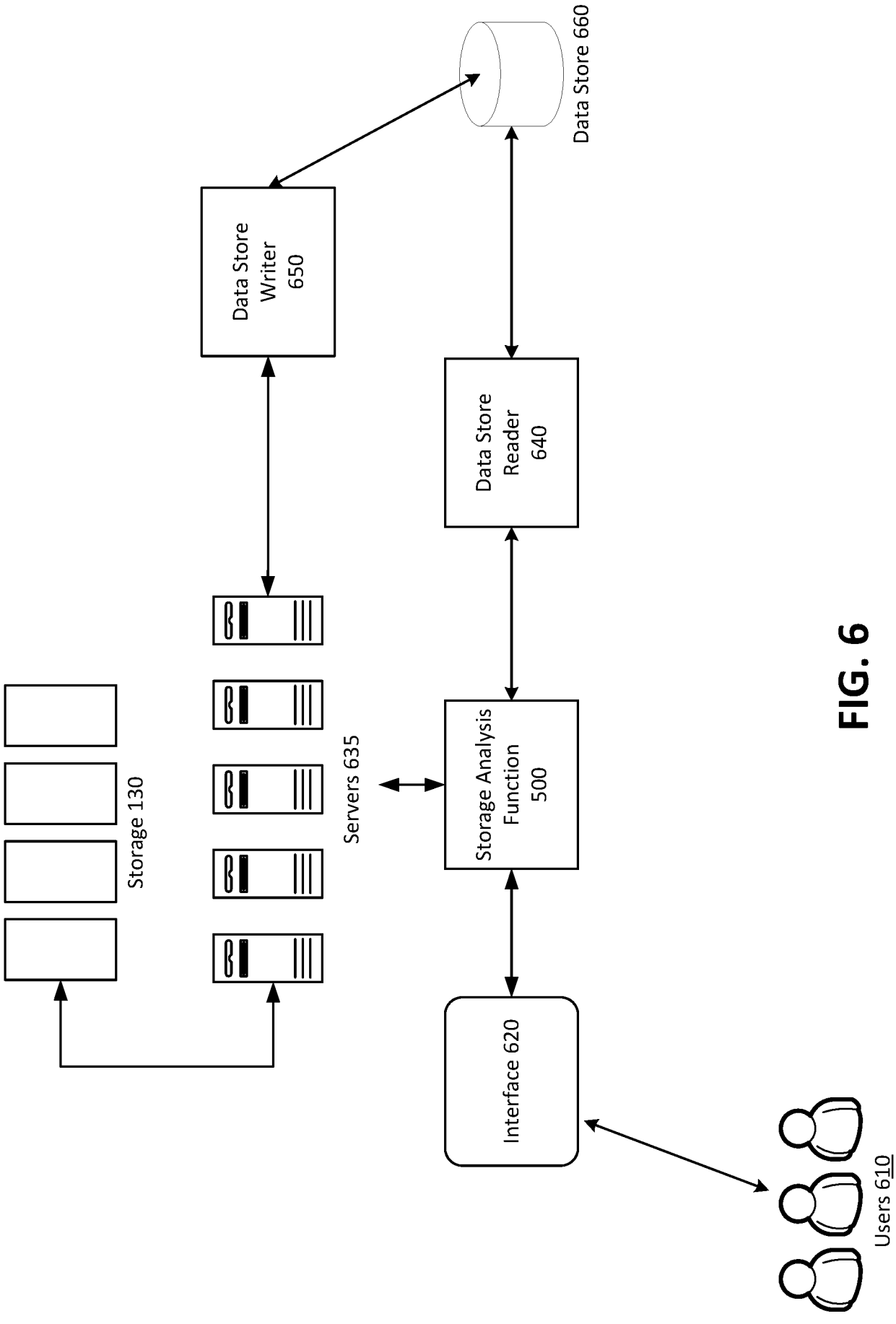

Data Store 660
Data Store Writer 650
Data Store Reader 640
Storage Analysis Function 500
Servers 635
Storage 130
Interface 620
Users 610

FIG. 6

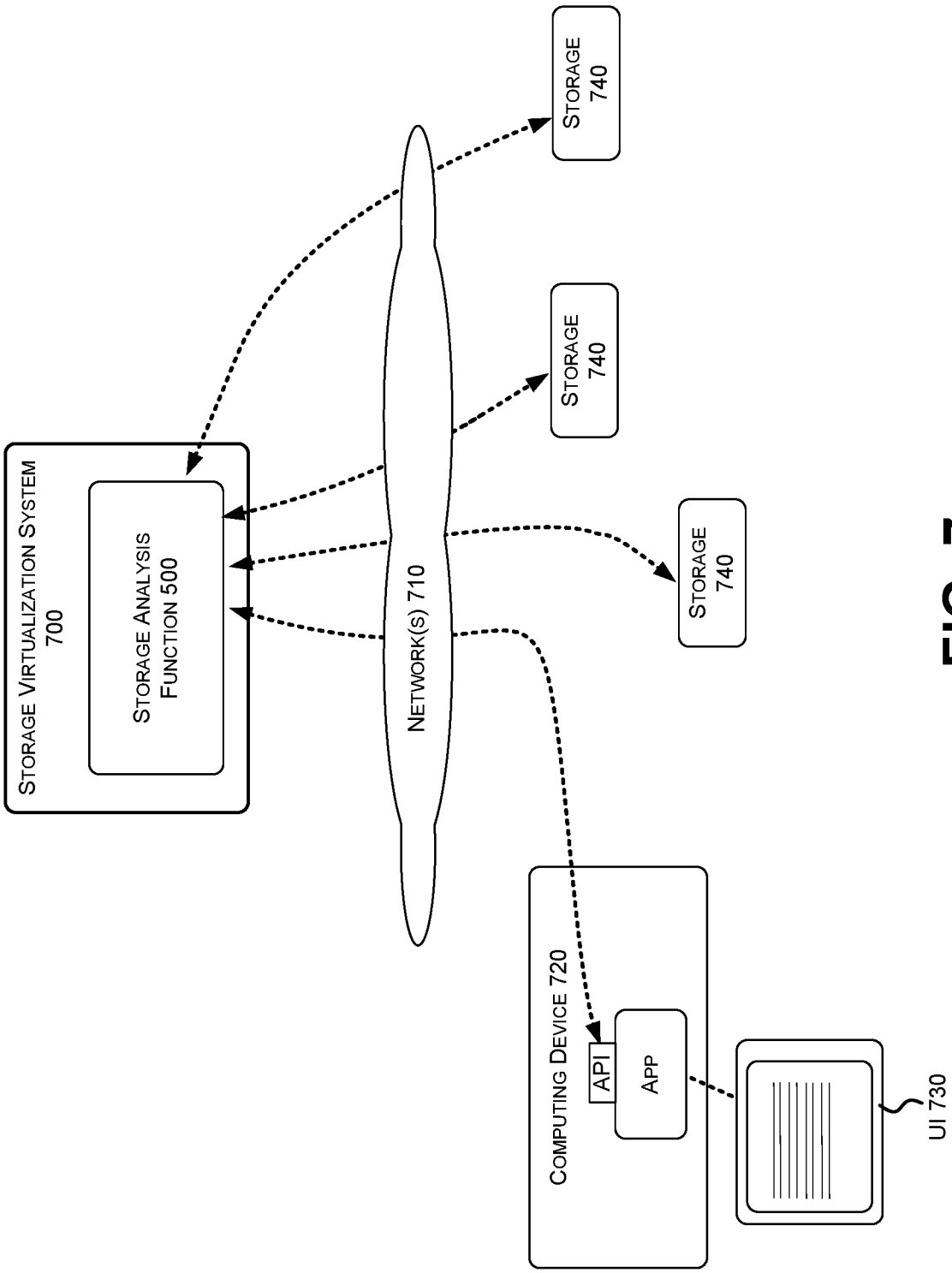

STORAGE VIRTUALIZATION SYSTEM 700
STORAGE ANALYSIS FUNCTION 500
NETWORK(S) 710
STORAGE 740
STORAGE 740
STORAGE 740
COMPUTING DEVICE 720
API
APP
UI 730

FIG. 7

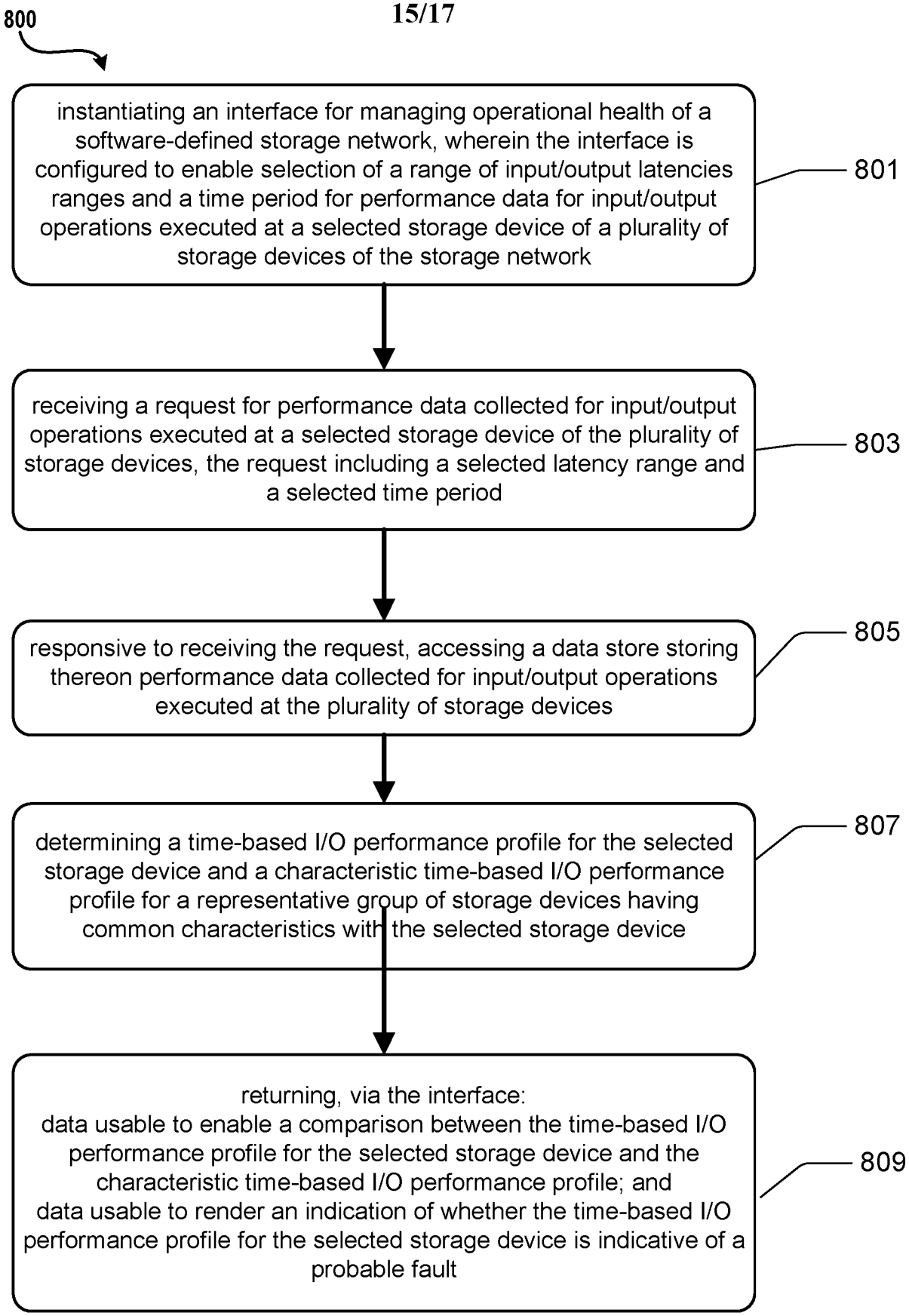

800
instantiating an interface for managing operational health of a software-defined storage network, wherein the interface is configured to enable selection of a range of input/output latencies ranges and a time period for performance data for input/output operations executed at a selected storage device of a plurality of storage devices of the storage network
801
receiving a request for performance data collected for input/output operations executed at a selected storage device of the plurality of storage devices, the request including a selected latency range and a selected time period
803
responsive to receiving the request, accessing a data store storing thereon performance data collected for input/output operations executed at the plurality of storage devices
805
determining a time-based I/O performance profile for the selected storage device and a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the selected storage device
807
returning, via the interface:
data usable to enable a comparison between the time-based I/O performance profile for the selected storage device and the characteristic time-based I/O performance profile; and
data usable to render an indication of whether the time-based I/O performance profile for the selected storage device is indicative of a probable fault
809

FIG. 8

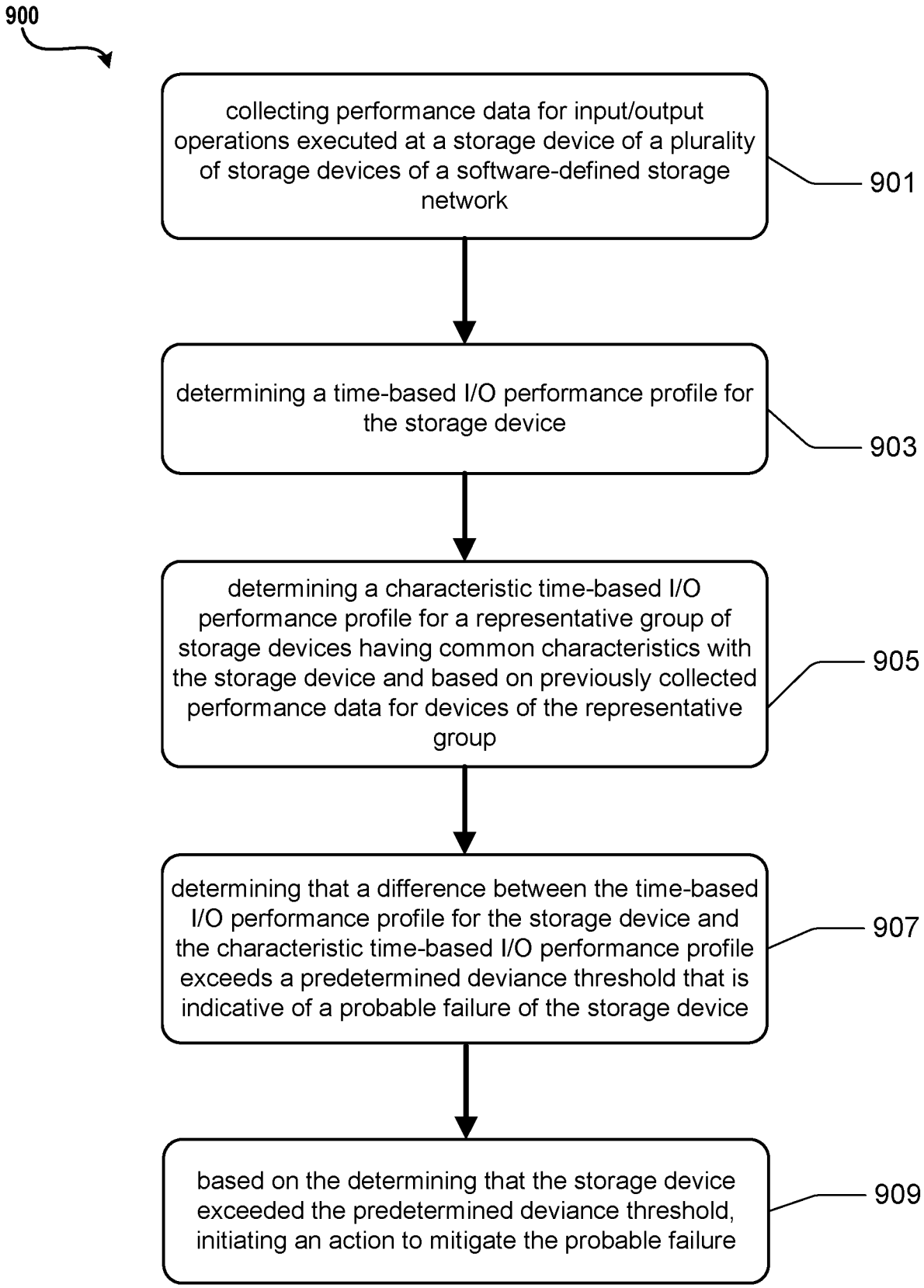

900
collecting performance data for input/output operations executed at a storage device of a plurality of storage devices of a software-defined storage network
901
determining a time-based I/O performance profile for the storage device
903
determining a characteristic time-based I/O performance profile for a representative group of storage devices having common characteristics with the storage device and based on previously collected performance data for devices of the representative group
905
determining that a difference between the time-based I/O performance profile for the storage device and the characteristic time-based I/O performance profile exceeds a predetermined deviance threshold that is indicative of a probable failure of the storage device
907
based on the determining that the storage device exceeded the predetermined deviance threshold, initiating an action to mitigate the probable failure
909

FIG. 9

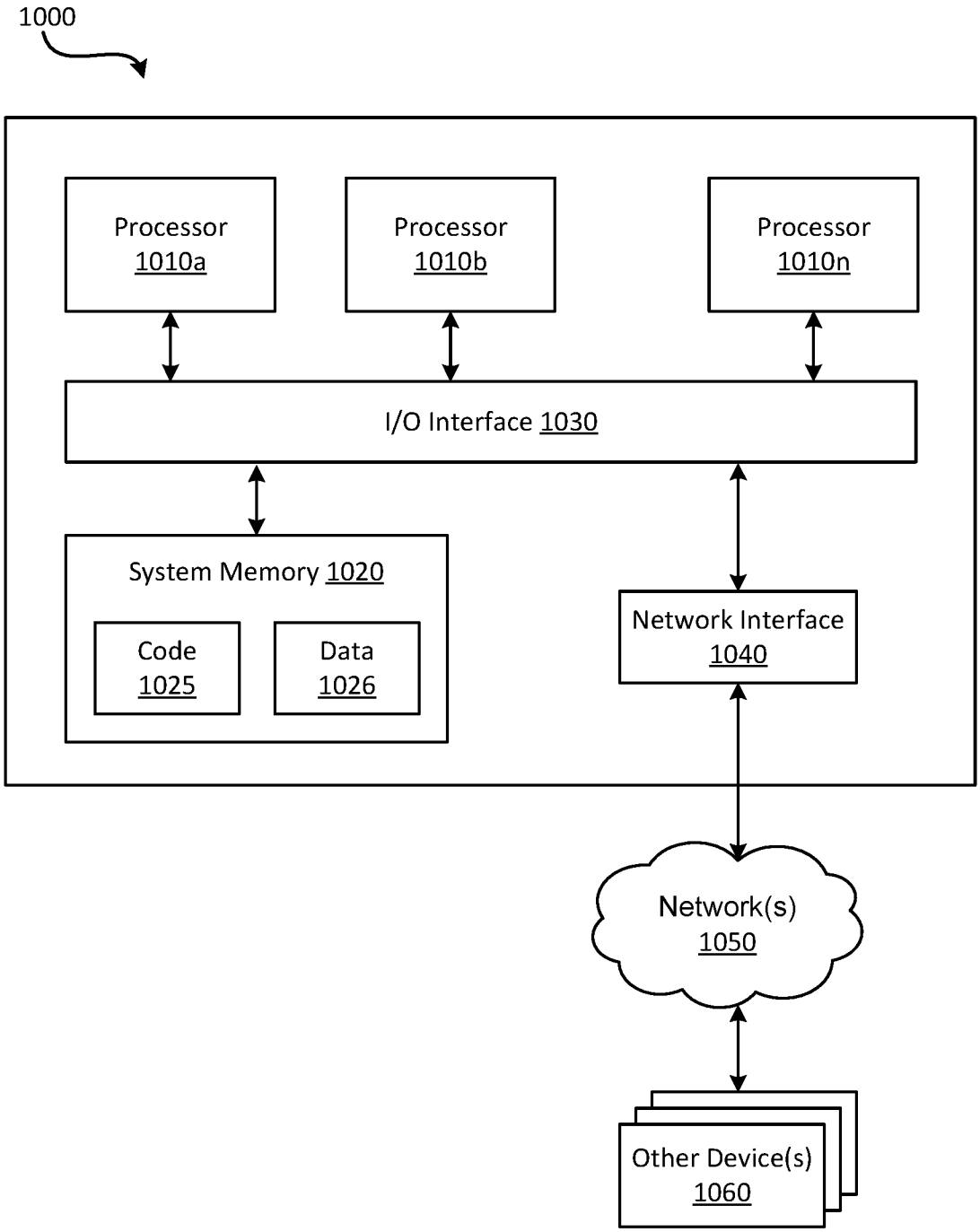

1000
Processor
1010a
Processor
1010b
Processor
1010n
I/O Interface 1030
System Memory 1020
Code
1025
Data
1026
Network Interface
1040
Network(s)
1050
Other Device(s)
1060

FIG. 10

# INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/015508

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F11/00 G06F11/07 G06F11/34 G06F11/32
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. |
|---|---|---|
| Y | US 2009/089503 A1 (YOSHIDA MASAHIRO [JP] ET AL) 2 April 2009 (2009-04-02)<br>abstract<br>paragraph [0007] - paragraph [0008]<br>paragraph [0034] - paragraph [0039]<br>paragraph [0053] - paragraph [0056]<br>----- | 1-15 |
| Y | US 2017/090779 A1 (BARZIK ZAH [IL] ET AL) 30 March 2017 (2017-03-30)<br>abstract<br>paragraph [0002] - paragraph [0004]<br>paragraph [0024]<br>paragraph [0036]<br>----- | 1-7,15 |

-/--

[X] Further documents are listed in the continuation of Box C.

[X] 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 | Date of mailing of the international search report |
|---|---|
| 8 May 2020 | 19/05/2020 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040,<br>Fax: (+31-70) 340-3016 | Lanchès, Philippe |

# INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/015508

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | US 2012/151276 A1 (BJORNER NIKOLAJ [US] ET AL) 14 June 2012 (2012-06-14)<br>abstract<br>paragraph [0008] - paragraph [0012]<br>paragraph [0025] - paragraph [0050]<br>paragraph [0063] - paragraph [0086]<br>----- | 8-14 |
| Y | US 2009/106602 A1 (PISZCZEK MICHAEL [US] ET AL) 23 April 2009 (2009-04-23)<br>abstract; figure 6<br>paragraph [0001] - paragraph [0002]<br>paragraph [0012]<br>paragraph [0022]<br>paragraph [0049] - paragraph [0060]<br>paragraph [0068]<br>paragraph [0076]<br>----- | 5-7 |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2020/015508

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| US 2009089503 A1 | 02-04-2009 | JP 4456626 B2 | 28-04-2010 |
| | | JP 2009086971 A | 23-04-2009 |
| | | US 2009089503 A1 | 02-04-2009 |
| US 2017090779 A1 | 30-03-2017 | NONE | |
| US 2012151276 A1 | 14-06-2012 | US 2012151276 A1 | 14-06-2012 |
| | | US 2014195860 A1 | 10-07-2014 |
| US 2009106602 A1 | 23-04-2009 | NONE | |