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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
09 January 2025 (09.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/008950 A1

(51) International Patent Classification:
H04L 67/1008 (2022.01) *H04W 24/02* (2009.01)
H04L 47/125 (2022.01)

(21) International Application Number:
PCT/IN2024/050931

(22) International Filing Date:
26 June 2024 (26.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
202321044339 03 July 2023 (03.07.2023) IN

(71) Applicant: **JIO PLATFORMS LIMITED** [IN/IN]; OFFICE-101, SAFFRON, NR. CENTRE POINT, PANCHWATI 5 RASTA, AMBAWADI, GUJARAT, AHMEDABAD 380006 (IN).

(72) Inventors: **BHATNAGAR, Aayush**; Tower-7, 15B, Beverly Park, Sector-14 Koper Khairane, Maharashtra, Navi Mumbai 400701 (IN). **BISHT, Birendra**; B-2101, Yashaskaram CHS, Plot -39, Sector -27, Kharghar, Maharashtra, Navi Mumbai 410210 (IN). **SINGH, Harbinder**; Wing B1, Flat No 402, Lakhani Suncoast, Sector 15, CBD Belapur, Maharashtra, Navi Mumbai 400614 (IN). **SINGH, Priyanka**; E-802 RiverScape CHS,Casa Rio, Palava City, Maharashtra, Dombivli East 421204 (IN). **AGGARWAL, Pravesh**; A-313, Raghubir Nagar, New Delhi, New Delhi 110027 (IN). **SOREN, Rohit**; Flat-106, HNo-84, Sultanpur, New Delhi, Delhi, New Delhi 110030 (IN). **SAHU, Bidhu**; 1702, E, RiverScape, CasaRio, Palava City, Dombivali East, Kalyan, Maharastra, Dombivali East 421204 (IN). **AMGOTHU, Ravi**; 1-111, Gopalapuram, A. kondurumandal, NTR district, Andhra Pradesh, Kondurumandal 521227 (IN). **DODDA, Sowmith**; H.no 1-104/T/1, near main road, Velikatta, Thorrur, Mahabubabad, Telangana, Mahabubabad 506163 (IN).

(74) Agent: **JAPHET, Chinthan**; K Law (Krishnamurthy and Co.), 4th Floor, Prestige Takt, No 23, Kasturba Road Cross, Bangalore 560 001 (IN).

(54) Title: METHOD AND SYSTEM FOR MANAGING A LOAD IN A NETWORK

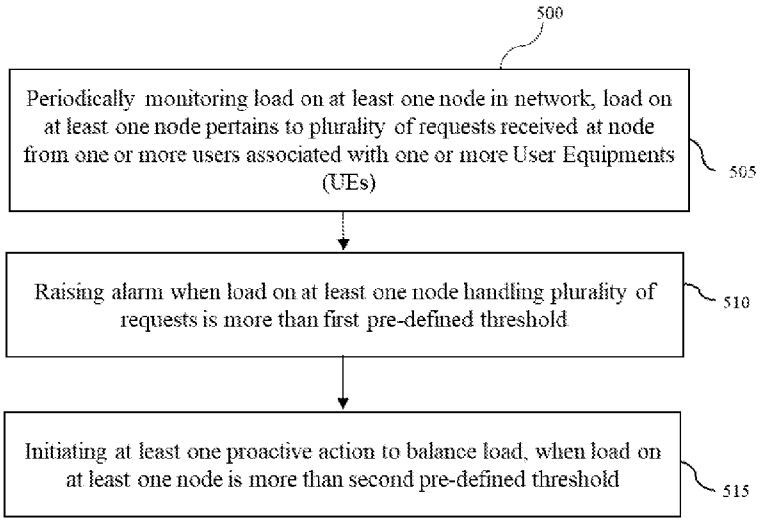

FIG. 5

(57) Abstract: The present disclosure relates to a system (120) and a method (500) for managing a load in a network (105) The system (120) includes a monitoring module (220) periodically monitor the load on at least one node (130) in the network (105), the load on the at least one node (130) pertains to a plurality of requests received at the at least one node (130) from one or more UEs (110). The system (120) includes an alarm module (225) to raise an alarm when the load on the at least one node (130) handling the plurality of requests is more than a first pre-defined threshold. The system (120) includes a load balancing module (230) initiate at least one proactive action to balance the load, when the load on the at least one node (130) is more than a second pre-defined threshold, thereby achieving optimum resource utilization and avoiding overload.

WO 2025/008950 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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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:

— *of inventorship (Rule 4.17(iv))*

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SYSTEM AND METHOD FOR MANAGING A LOAD IN A NETWORK

FIELD OF THE INVENTION

[0001] The present invention generally relates to wireless communication systems, and more particularly relates to a system and a method for managing a load in a network.

BACKGROUND OF THE INVENTION

[0002] Within a static pool and a dynamic pool of resources in a communication system, there may be scenarios wherein the static resource is stuck with a node and is not being utilised. The static pool of resources once assigned may not be available in the system.

[0003] In one scenario, when a load increases due to increase in the number of service requests such as a bunch of messages coming in at the same instance, or a bug etc., the pool of resources will get overloaded suddenly. This will slow down processing.

[0004] Also, during peak hours some nodes may get huge traffic disproportionate to the resources available for processing at that instance.

[0005] To improve efficiency of resource allocation in an application, a resource pool is used wherein a specific number of resources are pre-allocated during the startup of the application. When the application requires a resource, it can request one from the application pool, utilize it and subsequently return it after usage. However, in certain situations such as coding bug or unhandled events, the application may fail to free up resources after utilization. This can lead to resource leaks, where the resources are not properly released and eventually become depleted, potentially causing an outage.

[0006] The disadvantages associated are slow processing; resources will keep on getting overloaded and call failure. The incoming requests will keep on getting piled up and viciously because of the piling up of the load, more resources will be demanded, and the system may eventually fail.

[0007] It is desired that the allocation of resources is managed optimally in the system in such scenarios so that failures and outages are avoided.

[0008] Therefore, there is a need for an advancement for a system and method that can overcome at least one of the above shortcomings, particularly to manage resources in a communication network.

BRIEF SUMMARY OF THE INVENTION

[0009] One or more embodiments of the present disclosure provide a system and method for managing a load in a network.

[0010] In one aspect of the present invention, a system for managing a load in a network is disclosed. The system includes a monitoring module configured to periodically monitor the load on at least one node in the network. The load on the at least one node pertains to a plurality of requests received at the node from one or more users associated with one or more User Equipments (UEs). Each request is associated with availing one or more services. The system includes an alarm module configured to raise an alarm when the load on the at least one node handling the plurality of requests is more than a first pre-defined threshold. The system includes a load balancer module configured to initiate at least one proactive action to balance the load when the load on the at least one node is more than a second pre-defined threshold.

[0011] In one embodiment, the alarm raised by the alarm module is indicative of alerting network operators pertaining to the load at the at least one node. The load is more than the first pre-defined threshold.

[0012] In another embodiment, the at least one proactive action initiated by the load balancing module includes at least one of, restarting a current process thereby allowing a standby process to take over the current process, diverting the load to another node such as a standby node, changing status of the at least one node to a standby node and status of the standby node to an active node.

[0013] In yet another embodiment, the load balancer module is further configured to check in real time, the load on the at least one node when the at least one proactive action is initiated and halt the proactive action once the load on the at least one node is less than the second pre-defined threshold.

[0014] In yet another embodiment, the load on the at least one node is more than one of the first pre-defined threshold and the second pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage.

[0015] In yet another embodiment, the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by one of the one or more processors.

[0016] In yet another embodiment, an access module configured to continuously dump each type of resource usage data at pre-configured intervals into a log file.

[0017] In another aspect of the present invention, a method for managing a load in a network is disclosed. The method includes the steps of periodically monitoring load on at least one node in the network, the load on the at least one node pertains to a plurality of requests received at the node from one or more users associated with one or more User Equipments (UEs). Each request is associated with availing one or more services. The method further includes raising an alarm when the load on the at least one node handling the plurality of requests is more than a first pre-defined threshold. Further, the method includes initiating at least one proactive action to balance the load when the load on the at least one node is more than a second pre-defined threshold.

[0018] In another aspect of the present invention, a User Equipment (UE) includes one or more primary processors and a memory. The one or more primary processors is communicatively coupled to one or more processors and the memory. The memory stores instructions which when executed by the one or more primary processors causes

the UE to transmit one or more requests to the one or more processors in order to avail one or more services.

[0019] Other features and aspects of this invention will be apparent from the following description and the accompanying drawings. The features and advantages described in this summary and in the following detailed description are not all-inclusive, and particularly, many additional features and advantages will be apparent to one of ordinary skill in the relevant art, in view of the drawings, specification, and claims hereof. Moreover, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes and may not have been selected to delineate or circumscribe the inventive subject matter, resort to the claims being necessary to determine such inventive subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] The accompanying drawings, which are incorporated herein, and constitute a part of this disclosure, illustrate exemplary embodiments of the disclosed methods and systems in which like reference numerals refer to the same parts throughout the different drawings. Components in the drawings are not necessarily to scale, emphasis instead being placed upon clearly illustrating the principles of the present disclosure. Some drawings may indicate the components using block diagrams and may not represent the internal circuitry of each component. It will be appreciated by those skilled in the art that disclosure of such drawings includes disclosure of electrical components, electronic components or circuitry commonly used to implement such components.

[0021] **FIG. 1** is an exemplary block diagram of a communication system for managing a load in a network, according to one or more embodiments of the present disclosure;

[0022] **FIG. 2** is a block diagram of a system for managing the load in the network, according to one or more embodiments of the present disclosure;

[0023] **FIG. 3** is a schematic representation of the present system of **FIG. 2** workflow, according to one or more embodiments of the present disclosure;

[0024] **FIG. 4** shows an exemplary embodiment illustrating the system configured to receive a plurality of requests from one or more User Equipments (UEs), according to one or more embodiments of the present disclosure; and

[0025] **FIG. 5** shows a flow diagram of a method for managing a load in a network, according to one or more embodiments of the present disclosure.

[0026] The foregoing shall be more apparent from the following detailed description of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0027] Some embodiments of the present disclosure, illustrating all its features, will now be discussed in detail. It must also be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

[0028] Various modifications to the embodiment will be readily apparent to those skilled in the art and the generic principles herein may be applied to other embodiments. However, one of ordinary skill in the art will readily recognize that the present disclosure including the definitions listed here below are not intended to be limited to the embodiments illustrated but is to be accorded the widest scope consistent with the principles and features described herein.

[0029] A person of ordinary skill in the art will readily ascertain that the illustrated steps detailed in the figures and here below are set out to explain the exemplary embodiments shown, and it should be anticipated that ongoing technological development will change the manner in which particular functions are performed. These examples are presented herein for purposes of illustration, and not limitation. Further, the boundaries of the functional building blocks have been arbitrarily defined

herein for the convenience of the description. Alternative boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed. Alternatives (including equivalents, extensions, variations, deviations, etc., of those described herein) will be apparent to persons skilled in the relevant art(s) based on the teachings contained herein. Such alternatives fall within the scope and spirit of the disclosed embodiments.

[0030] As per various embodiments depicted, the present invention discloses the system and method for managing a load in a network. The load is managed optimally in the system and method so that failures and outages are avoided.

[0031] Referring to **FIG. 1**, **FIG. 1** illustrates an exemplary block diagram of a communication system **100** for managing a load in a network **105,** according to one or more embodiments of the present disclosure. The communication system **100** includes one or more User Equipment's (UEs) **110**. Each of the at least one UE **110** from the one or more UEs **110a, 110b, ...110n** is configured to connect to a server **115** via the network **105**. For the purpose of description and explanation, the description will be explained with respect to one or more UEs **110**, or to be more specific will be explained with respect to a first UE **110a**, a second UE **110b**, and a third UE **110c**, and should nowhere be construed as limiting the scope of the present disclosure.

[0032] In an embodiment, each of the first UE **110a**, the second UE **110b**, and the third UE **110c** is one of, but are not limited to, any electrical, electronic, electro-mechanical or an equipment and a combination of one or more of the above devices such as virtual reality (VR) devices, augmented reality (AR) devices, laptop, a general-purpose computer, desktop, personal digital assistant, tablet computer, mainframe computer, or any other computing device. Each of the first UE **110a**, the second UE **110b**, and the third UE **110c** is configured to transmit one or more requests from the one or more UEs **110** to one or more processors **202** (as shown in **FIG. 2**) to avail one or more services. In one embodiment, the one or more services include, but not limited to, messaging services, call services, and streaming services.

[0033] The server **115** may include by way of example but not limitation, one or more of a standalone server, a server blade, a server rack, a bank of servers, a server farm, hardware supporting a part of a cloud service or system, a home server, hardware running a virtualized server, one or more processors executing code to function as a server, one or more machines performing server-side functionality as described herein, at least a portion of any of the above, some combination thereof. In an embodiment, the entity may include, but is not limited to, a vendor, a network operator, a company, an organization, a university, a lab facility, a business enterprise side, a defense facility side, or any other facility that provides service.

[0034] The network **105** includes, by way of example but not limitation, one or more of a wireless network, a wired network, an internet, an intranet, a public network, a private network, a packet-switched network, a circuit-switched network, an ad hoc network, an infrastructure network, a Public-Switched Telephone Network (PSTN), a cable network, a cellular network, a satellite network, a fiber optic network, or some combination thereof. The network **105** may include, but is not limited to, a Third Generation (3G), a Fourth Generation (4G), a Fifth Generation (5G), a Sixth Generation (6G), a New Radio (NR), a Narrow Band Internet of Things (NB-IoT), an Open Radio Access Network (O-RAN), and the like.

[0035] The network **105** may also include, by way of example but not limitation, at least a portion of one or more networks having one or more nodes that transmit, receive, forward, generate, buffer, store, route, switch, process, or a combination thereof, etc. one or more messages, packets, signals, waves, voltage or current levels, some combination thereof, or so forth. The network may also include, by way of example but not limitation, one or more of a wireless network, a wired network, an internet, an intranet, a public network, a private network, a packet-switched network, a circuit-switched network, an ad hoc network, an infrastructure network, a Public-Switched Telephone Network (PSTN), a cable network, a cellular network, a satellite network, a fiber optic network, a VOIP or some combination thereof.

[0036] The communication system **100** further includes an at least one node **130**. In one embodiment, the communication system **100** further includes one or more nodes**.** In another embodiment, the one or more nodes include, but not limited to, an active node **140** and a standby node **150**. The at least one node **130** is communicably coupled to the server **115** via the network **105**. The at least one node **130** is connected to send, receive, or forward data to the server **115** within the network **105**. The active node **140** actively performs tasks and processing requests, and handling a workload. The active node **140** is responsible for processing user requests, running applications, or managing services. The standby node **150** is not actively processing user requests or handling the primary workload. Instead, the standby node **150** remains in a standby state, ready to take over operations in the event of a failure or when a predefined threshold triggers a failover. The standby node **150** essentially serves as a backup to ensure continuity and minimize downtime.

[0037] The communication system **100** further includes a system **120** communicably coupled to the server **115** and each of the first UE **110a**, the second UE **110b**, and the third UE **110c** via the network **105**. The system **120** is adapted to be embedded within the server **115** or is embedded as the individual entity. However, for the purpose of description, the system **120** is described as an integral part of the server **115**, without deviating from the scope of the present disclosure.

[0038] Operational and construction features of the system **120** will be explained in detail with respect to the following figures.

[0039] Referring to **FIG. 2**, **FIG. 2** illustrates a block diagram of the system **120** for managing the load in the network **105**, according to one or more embodiments of the present disclosure. The system **120** includes one or more processors **202**, a memory **204**, an input/output interface unit **206**, a display **208**, and an input device **210**. The one or more processor **202**, hereinafter referred to as the processor **202** may be implemented as one or more microprocessors, microcomputers, microcontrollers, digital signal processors, central processing units, state machines, logic circuitries,

single board computers, and/or any devices that manipulate signals based on operational instructions. As per the illustrated embodiment, the system **120** includes one processor **202**. However, it is to be noted that the system **120** may include multiple processors as per the requirement and without deviating from the scope of the present disclosure.

[0040] The information related to the load pertains to a plurality of requests received on at least one node **130** from the one or more UEs **110** may be provided or stored in the memory **204**. Among other capabilities, the processor **202** is configured to fetch and execute computer-readable instructions stored in the memory **204**. The memory **204** may be configured to store one or more computer-readable instructions or routines in a non-transitory computer-readable storage medium, which may be fetched and executed to create or share data packets over a network service. The memory **204** may include any non-transitory storage device including, for example, volatile memory such as RAM, or non-volatile memory such as EPROMs, FLASH memory, unalterable memory, and the like.

[0041] The information related to the load pertains to the plurality of requests received at the at least one node **130** from the one or more UEs **110** may further be configured to render on the I/O interface unit **206**. The I/O interface unit **206** may include functionality similar to at least a portion of functionality implemented by one or more computer system interfaces such as those described herein and/or generally known to one having ordinary skill in the art. The I/O interface unit **206** may be rendered on a display **208**, implemented using an LCD display technology, an OLED display technology, and/or other types of conventional display technology. The display **208** may be integrated within the system **120** or connected externally. Further the input device(s) **210** may include, but not limited to, keyboard, buttons, scroll wheels, cursors, touchscreen sensors, audio command interfaces, magnetic strip reader, optical scanner, etc.

[0042] The database **240** is configured to store the plurality of requests made by the UE **110** over time. Further, the database **240** provides structured storage, support for complex queries, and enables efficient data retrieval and analysis. The database **240** is one of, but is not limited to, one of a centralized database, a cloud-based database, a commercial database, an open-source database, a distributed database, an end-user database, a graphical database, a No-Structured Query Language (NoSQL) database, an object-oriented database, a personal database, an in-memory database, a document-based database, a time series database, a wide column database, a key value database, a search database, a cache databases, and so forth. The foregoing examples of database **240** types are non-limiting and may not be mutually exclusive e.g., a database can be both commercial and cloud-based, or both relational and open-source, etc.

[0043] Further, the processor **202**, in an embodiment, may be implemented as a combination of hardware and programming (for example, programmable instructions) to implement one or more functionalities of the processor **202**. In the examples described herein, such combinations of hardware and programming may be implemented in several different ways. For example, the programming for the processor **202** may be processor-executable instructions stored on a non-transitory machine-readable storage medium and the hardware for processor **202** may comprise a processing resource (for example, one or more processors), to execute such instructions. In the present examples, the memory **204** may store instructions that, when executed by the processing resource, implement the processor **202**. In such examples, the system **120** may comprise the memory **204** storing the instructions and the processing resource to execute the instructions, or the memory **204** may be separate but accessible to the system **120** and the processing resource. In other examples, the processor **202** may be implemented by electronic circuitry.

[0044] In order for the system **120** to manage the load in the network **105**, the processor **202** includes a monitoring module **220**, an alarm module **225**, a load balancing module **230**, and an access module **235** communicably coupled to each other for managing the load in the network **105**.

[0045] The monitoring module **220**, the alarm module **225**, the load balancing module **230**, and the access module **235** in an embodiment, may be implemented as a combination of hardware and programming (for example, programmable instructions) to implement one or more functionalities of the processor **202**. In the examples described herein, such combinations of hardware and programming may be implemented in several different ways. For example, the programming for the processor **202** may be processor-executable instructions stored on a non-transitory machine-readable storage medium and the hardware for the processor may comprise a processing resource (for example, one or more processors), to execute such instructions. In the present examples, the memory **204** may store instructions that, when executed by the processing resource, implement the processor. In such examples, the system **120** may comprise the memory **204** storing the instructions and the processing resource to execute the instructions, or the memory **204** may be separate but accessible to the system **120** and the processing resource. In other examples, the processor **202** may be implemented by electronic circuitry.

[0046] The monitoring module **220** is configured to periodically monitor the load on an at least one node **130** in the network **105**. The load on the at least one node **130** pertains to a plurality of requests received at the node from one or more users associated with one or more User Equipments (UEs) **110**. In an embodiment, each request is associated with availing one or more services. In one embodiment, the one or more services include, but not limited to, messaging services, call services, and streaming services. When the load on the at least one node **130** handling the plurality of requests is more than a first pre-defined threshold, the alarm module **225** raises an alarm to indicate network operators pertaining to the load on the at least one node **130**. In an embodiment, the load includes, but is not limited to the plurality of requests and packet processing.

[0047] The alarm module **225** is configured to raise the alarm when the load on the at least one node **130** handling the plurality of requests is more than the first pre-defined threshold. The first pre-defined threshold is defined by the network operators

based on the received plurality of requests performed by the at least one node **130** within a predetermined interval. In an embodiment, the alarm raised by the alarm module **225** is indicative of alerting network operators pertaining to the load at the at least one node **130** when the load is more than the first pre-defined threshold, thereby facilitating the network operators to resolve the load on the at least one node **130**. In an embodiment, the load on the at least one node **130** is more than one of the first pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage.

[0048] In an exemplary embodiment, the load monitoring enables early detection of faults or anomalies that indicates hardware failures, software errors, or network congestion in the network **105**. When the load exceeds the second predefined threshold as defined by the network operator, the alerts are triggered. Owing to the triggering, the network operators are notified, and accordingly facilitating rapid troubleshooting and resolution of issues. Further, the load monitoring aids in tracking resource usage patterns. Based on this, the network operators identify when the resources are approaching capacity limits. Owing to the load monitoring, the network operators are configured to initiate the at least one proactive action to optimize resource allocation or scale resources as per requirement. In one embodiment, the first predefined threshold is pre-defined or dynamically set by one of the one or more processors **202**. When the load has exceeded the first pre-defined threshold, the load balancing module **230** initiates at least one proactive action to balance the load.

[0049] The load balancing module **230** is configured to initiate at least one proactive action to balance the load when the load on the at least one node **130** is more than a second pre-defined threshold. The second pre-defined threshold is defined by the network operators based on the received plurality of requests performed by the at least one node **130** within the predetermined interval. In an embodiment, the at least one proactive action initiated by the load balancing module **230** includes at least one of, but not limited to, restarting a current process thereby allowing the standby node **150** to take over the current process, diverting the load from the active node **140** to the

standby node **150**, changing status of the at least one node **130** to the standby node **150** and the status of the standby node **150** to the active node **140**. In an embodiment, the load on the at least one node **130** is more than one of the second pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage.

[0050] Further, the load balancing module **230** is configured to check the load on the at least one node **130** in real time when the at least one proactive action is initiated. Once the load on the at least one node **130** is less than the second predefined threshold, the load balancing module **230** is configured to halt the proactive action. The halt process refers to an action taken by the load balancing module **230** to stop or pause the at least one proactive action, when the load on the at least one node **130** is below the second predefined threshold. More specifically, as the monitoring module **220** detects that the load on the at least one node **130** is lesser than the second predefined threshold, the load balancing module **230** triggers the halt process to prevent unnecessary utilization of resources or potential overload on another node. This ensures that the system **120** maintains stability and optimal performance by dynamically adjusting the at least one proactive action based on real-time load conditions.

[0051] In an embodiment, the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by one of the one or more processors **202**. In another embodiment, the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by the network operators based on the received plurality of requests performed by the at least one node **130** within the predetermined interval.

[0052] Upon halting the proactive action, the load on the at least one node **130** is less than the second pre-defined threshold. The access module **235** is configured to continuously dump each type of resource usage data at pre-configured intervals into a log file, thereby providing developers access to historical records of resource allocation and utilization patterns. By doing so, the system **120** is configured to

achieve optimum resource utilization, no need of manual intervention when maximum capacity of the plurality of requests is breached as in case of the second pre-defined threshold, and the proactive action is taken as described above so as to reduce processing time, and avoiding overloading of the plurality of requests which improves overall system performance.

[0053] **FIG. 3** is a schematic representation of the system **120** in which various entities operations are explained, according to one or more embodiments of the present disclosure. Referring to **FIG. 3**, describes the system **120** for managing the load in the network **105**. It is to be noted that the embodiment with respect to **FIG. 3** will be explained with respect to the first UE **110a** for the purpose of description and illustration and should nowhere be construed as limited to the scope of the present disclosure.

[0054] As mentioned earlier in **FIG.1**, each of the first UE **110a,** the second UE **110b**, and the third UE **110c** may include an external storage device, a bus, a main memory, a read-only memory, a mass storage device, communication port(s), and a processor. The exemplary embodiment as illustrated in the **FIG. 3** will be explained with respect to the first UE **110a**. The first UE **110a** includes one or more primary processors **305** communicably coupled to the one or more processors **205** of the system **125**.

[0055] The one or more primary processors **305** are coupled with a memory **310** storing instructions which are executed by the one or more primary processors **305**. Execution of the stored instructions by the one or more primary processors **305** enables the first UE **110a** to transmit the one or more requests to the one or more processors in order to avail one or more services.

[0056] As mentioned earlier in **FIG.2**, the one or more processors **202** of the system **120** is configured to perform monitoring the load, raising the alarm when the when the load on the at least one node **130** handling plurality of requests is more than the first

pre-defined threshold, and initiating the proactive action to balance the load when the load on the at least one node **130** is more than the second pre-defined threshold.

[0057] As per the illustrated embodiment, the system **120** includes the one or more processors **202**, the memory **204**, the input/output interface unit **206**, the display **208**, and the input device **210**. The operations and functions of the one or more processors **202**, the memory **204**, the input/output interface unit **206**, the display **208**, and the input device **210** are already explained in **FIG. 2**. For the sake of brevity, a similar description related to the working and operation of the system **120** as illustrated in **FIG. 2** has been omitted to avoid repetition.

[0058] Further, the processor **202** includes the monitoring module **220**, the alarm module **225**, the load balancing module **230,** and the access module **235**. The operations and functions of the monitoring module **220**, the alarm module **225**, the load balancing module **230,** and the access module **235** are already explained in **FIG. 2**. Hence, for the sake of brevity, a similar description related to the working and operation of the system **120** as illustrated in **FIG. 2** has been omitted to avoid repetition. The limited description provided for the system **120** in **FIG. 3**, should be read with the description as provided for the system **120** in the **FIG. 2** above, and should not be construed as limiting the scope of the present disclosure.

[0059] **FIG. 4** shows an exemplary embodiment illustrating the system **120** configured to receive a plurality of requests from one or more User Equipments (UEs) **110**, in accordance with the present disclosure.

[0060] At step **402**, receiving the plurality of requests from one or more users associated with the one or more UEs **110** to the one or more processors **202** of the system **120**. The system **120** is configured to handle the plurality of requests (for example 500 requests) in a given time. When the system **120** receives the plurality of requests which causes multiple errors such as network traffics, bugs or sudden surge in the servicing requests. For avoiding the errors, the system **120** needs to monitor the

plurality of requests received at the node. The system **120** is configured to utilize certain predefined thresholds to avoid the errors.

[0061] At step **404**, the monitoring module is configured for periodically monitoring the load on at least one node **130** pertains to the plurality of requests (for example 5000) received at the node from the one or more users associated with one or more User Equipments (UEs).

[0062] At step **406,** the system **120** is configured for handling 500 requests in a given time but the system **120** gets 1000 requests. The system **120** may have two pre-defined thresholds. In one embodiment, the pre-defined thresholds may include, but not limited to, the first pre-defined threshold, and the second pre-defined threshold. In another embodiment, for example, the first pre-defined threshold may be 400 requests and the second pre-defined threshold may be 500 requests. The alarm module **225** is configured to raise the alarm to alert the network operators to avoid overload of the plurality of requests when the system **120** reaches the first pre-defined threshold (such as 400 requests).

[0063] At step **408**, the load balancing module **230** initiates the proactive action such as but not limited to, restarting the current process and diverting the load to another node such as the standby node **150**, changing status of the at least one node to the standby node **150** and status of the standby node **150** to an active node **140**. In one embodiment, both the first and second pre-defined thresholds are breached, there might be the coding bug and the resource might be leaking very quickly. In this scenario, the node will reboot and enter an auto recovery process and divert the load from the active node **140** to the standby node **150** if available.

[0064] At step **410**, continuously dumping each type of load usage data at pre-configured intervals into the log file by the access module **235**, thereby providing developers access to historical records of load allocation and utilization patterns in the database **240**.

[0065] **FIG. 5** is a flow chart of the method **500** for managing the load in the network **105**, according to one or more embodiments of the present invention. For the purpose of description, the method **500** is described with the embodiments as illustrated in **FIG. 2** and should nowhere be construed as limiting the scope of the present disclosure.

[0066] At step **505,** the method **500** includes the step of periodically monitoring, the load on at least one node **130** in the network **105** by the monitoring module **220**. The load on the at least one node **130** pertains to the plurality of requests received at the node from one or more users associated with one or more User Equipments (UEs) **110**. In an embodiment, each request is associated with availing one or more services. The one or more services include, but not limited to, transmitting the request, and monitoring the load via the network **105**. When the load on the at least one node **130** handling the plurality of requests is more than the first pre-defined threshold, then the alarm module **225** raises an alarm to indicate network operators pertaining to the load on the at least one node **130**.

[0067] At step **510**, the method **500** includes the step of raising the alarm by the alarm module **225** when the load on the at least one node **130** handling the plurality of requests is more than the first pre-defined threshold. The first pre-defined threshold is defined by the network operators based on the received plurality of requests performed by the at least one node **130** within a predetermined interval.

[0068] In an embodiment, the alarm raised by the alarm module **225** is indicative of alerting network operators pertaining to the load at the at least one node **130** when the load is more than the first pre-defined threshold, thereby facilitating the network operators to resolve the load on the at least one node **130**. In an embodiment, the load on the at least one node **130** is more than one of the first pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage. In one embodiment, the first predefined threshold is pre-defined or dynamically set by one of the one or more processors **202**. When the load has exceeded

the first pre-defined threshold, the load balancing module **230** initiates at least one proactive action to balance the load.

[0069] At step **515**, the method **500** includes the step of initiating at least one proactive action to balance the load by using the load balancing module **230** when the load on the at least one node **130** is more than a second pre-defined threshold. In an embodiment, the at least one proactive action initiated by the load balancing module **230** includes at least one of, restarting a current process thereby allowing a standby process to take over the current process, diverting the load from an active node **140** to a standby node **150**, changing status of the at least one node **130** to the standby node **150** and status of the standby node **150** to the active node **140**. In an embodiment, the load on the at least one node **130** is more than one of the second pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage.

[0070] In an exemplary embodiment, the at least one node **130** is configured to receive the plurality of requests from the UE 110. The network operator defines the first and the second predefined thresholds corresponding to the load on the at least one node **130** as 90 and 100 requests, respectively. During operation, the monitoring module **220** is configured to monitor the load at the at least one node **130**. At an instance of time let us consider, the at least one node receives 91 requests, which is greater than the first predefined threshold. Accordingly, the alarm module **225** raises the alarm to indicate that the load is greater than the first predefined threshold.

[0071] Further, the monitoring module **220** continues to monitor the at least one node **130**. At another instance of time, let us consider the at least one node **130** receives 102 requests, which are greater than the second predefined threshold. In response, the load balancing module **230** is configured to initiate the at least one proactive action. One example of the at least one proactive action is allowing the standby process to take over the current process running in the at least one node **130**. In order to do so, the load balancing module **230** is configured to restart the

current process. As a result, the plurality of requests is queued in the standby node **150**.

[0072] Another example of the at least one proactive action includes diverting the load received at the at least one node **130** to the standby node **150** until the load at the at least one node **130** is lesser than the first predefined threshold. Yet another example of the at least one proactive action includes changing status of the at least one node **130** to indicate the at least one node **130** is the standby node **150**, and the status of the standby node **150** to indicate the standby node **150** is the active node **140**.

[0073] Further, the load balancing module **230** is further configured to check in real time the load on the at least one node **130** when the at least one proactive action is initiated. Once the load on the at least one node **130** is less than the second pre-defined threshold, the load balancing module **230** is configured to halt the proactive action. The halt process refers to an action taken by the load balancing module **230** to stop or pause the at least one proactive action, when the load on the at least one node **130** is below the second predefined threshold. More specifically, as the monitoring module **220** detects that the load on the at least one node **130** is lesser than the second predefined threshold, the load balancing module **230** triggers the halt process to prevent unnecessary utilization of resources or potential overload on other nodes. This ensures that the system **120** maintains stability and optimal performance by dynamically adjusting the proactive actions based on real-time load conditions.

[0074] In an embodiment, the second predefined threshold is pre-defined or dynamically set by one of the one or more processors **202**. In another embodiment, the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by one of the one or more processors **202**. In yet another embodiment, the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by the network operators based on the received plurality of requests performed by the at least one node **130** within the predetermined interval. By doing so, the system **120** is configured to achieve optimum resource utilization, no need of

manual intervention when maximum capacity of the plurality of requests is breached as in case of the second pre-defined threshold, and the proactive action is taken as described above so as to reduce processing time, and avoiding overloading of the plurality of requests which improves overall system performance.

[0075] The present invention further discloses a non-transitory computer-readable medium having stored thereon computer-readable instructions. The computer-readable instructions are executed by a processor **205**. The processor **205** is configured to periodically monitor the load on at least one node **130** in the network, the load on the at least one node **130** pertains to the plurality of requests received at the node from one or more users associated with one or more User Equipments (UEs) **110**, each request associated with availing one or more services. The processor **205** is configured to raise the alarm when the load on the at least one node **130** handling the plurality of requests is more than the first pre-defined threshold. Further, the processor **205** is configured to initiate at least one proactive action to balance the load, when the load on the at least one node **130** is more than the second pre-defined threshold.

[0076] A person of ordinary skill in the art will readily ascertain that the illustrated embodiments and steps in description and drawings (FIG.1-5) are set out to explain the exemplary embodiments shown, and it should be anticipated that ongoing technological development will change the manner in which particular functions are performed. These examples are presented herein for purposes of illustration, and not limitation. Further, the boundaries of the functional building blocks have been arbitrarily defined herein for the convenience of the description. Alternative boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed. Alternatives (including equivalents, extensions, variations, deviations, etc., of those described herein) will be apparent to persons skilled in the relevant art(s) based on the teachings contained herein. Such alternatives fall within the scope and spirit of the disclosed embodiments.

[0077] The present disclosure incorporates technical advancement of proactively monitoring and auditing the load, raising an alarm for manual intervention to avoid the overloading of the plurality of requests when the first pre-defined threshold is reached. Further the invention provides the proactive actions for the same such as restarting the process, diverting the load to a standby turned active node, changing from the active node to the standby node by restarting and changing the standby node to the active node to balance the load when the second pre-defined threshold is reached. The present disclosure collects detailed information about the specific resources that are being exhausted and dumps the same for post recovery debugging process. Additionally, as part of the recovery mechanism, it automatically restarts the affected process, ensuring continuous availability of the resource and improving an overall system performance.

[0078] The present disclosure significantly achieves optimum resource utilization. When maximum capacity of the requests is breached as in case of second predefined threshold, an automatic action is taken without manual intervention. By doing so the system and method to improve overall system performance and ensure continuous availability of the resources. Further, the load is managed optimally in the system in such scenarios so that failures and outages are avoided.

[0079] The present invention offers multiple advantages over the prior art and the above listed are a few examples to emphasize on some of the advantageous features. The listed advantages are to be read in a non-limiting manner.

REFERENCE NUMERALS

[0080] Communication system - 100;

[0081] Network - 105;

[0082] User Equipment - 110;

[0083] Server - 115;

[0084] System - 120;

[0085] At least one node- 130;

[0086] Active node - 140;

[0087] Standby Node – 150;

[0088] One or more processor -202;

[0089] Memory – 204;

[0090] Input/Output interface unit – 206;

[0091] Display- 208;

[0092] Input device -210;

[0093] Monitoring Module- 220;

[0094] Alarm Module - 225;

[0095] Load Balancing Module - 230;

[0096] Access Module - 235;

[0097] Database - 240;

[0098] One or more primary processors – 305;

[0099] Memory of user equipment – 310.

We Claim:

1. A method (500) for managing a load in a network (105), the method (500) comprises the steps of:

periodically monitoring (505), by one or more processors (202), load on at least one node (130) in the network (105), the load on the at least one node (130) pertains to a plurality of requests received at the at least one node (130) from one or more users associated with one or more User Equipments (UEs), each request associated with availing one or more services;

raising, by the one or more processors (202), an alarm when the load on the at least one node (130) handling the plurality of requests is more than a first pre-defined threshold; and

initiating, by the one or more processors (202), at least one proactive action to balance the load, when the load on the at least one node (130) is more than a second pre-defined threshold.

2. The method (500) as claimed in claim 1, wherein the alarm raised is indicative of alerting network operators pertaining to the load at the at least one node (130), the load is more than the first pre-defined threshold.

3. The method (500) as claimed in claim 1, wherein the at least one proactive action initiated includes at least one of, restarting a current process thereby allowing a standby process to take over the current process, diverting the load to another node such as a standby node (150), changing status of the at least one node to the standby node (150) and status of the standby node (150) to an active node (140).

4. The method (500) as claimed in claim 1, wherein the step of, initiating, at least one proactive action to balance the load, when the load on the at least one node (130) is more than the second pre-defined threshold, further includes the steps of:

checking in real time, by the one or more processors (202), the load on the at least one node (130) when the at least one proactive action is initiated; and

halting, by the one or more processors (202), the proactive action once the load on the at least one node (130) is less than the second pre-defined threshold.

5. The method (500) as claimed in claim 1, wherein the load on the at least one node (130) is more than one of the first pre-defined threshold and the second pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage.

6. The method (500) as claimed in claim 1, wherein the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by one of one or more processors (202).

7. The method (500) as claimed in claim 1, wherein the method (500) further comprises the step of:

continuous dumping, by the one or more processors (202), each type of resource usage data at pre-configured intervals into a log file,

8. A system (120) for managing a load in a network (105), the system (120) comprising:

a monitoring module (220) configured to, periodically monitor, load on at least one node (130) in the network (105), the load on the at least one node (130) pertains to a plurality of requests received at the at least one node (130) from one or more users associated with one or more User Equipments (UEs), each request associated with availing one or more services;

an alarm module (225) configured to, raise, an alarm when the load on the at least one node (130) handling the plurality of requests is more than a first pre-defined threshold; and

a load balancing module (230) configured to, initiate, at least one proactive action to balance the load, when the load on the at least one node (130) is more than a second pre-defined threshold.

9. The system (120) as claimed in claim 8, wherein the alarm raised by the alarm module (225) is indicative of alerting network operators pertaining to the load at the at least one node (130), the load is more than the first pre-defined threshold.

10. The system (120) as claimed in claim 8, wherein the at least one proactive action initiated by the load balancing module (235) includes at least one of, restarting a current process thereby allowing a standby process to take over the current process, diverting the load to another node such as a standby node (150), changing status of the at least one node to the standby node (150) and status of the standby node (150) to an active node (140).

11. The system (120) as claimed in claim 8, wherein on initiating, at least one proactive action to balance the load, when the load on the at least one node (130) is more than the second pre-defined threshold, the load balancing module (235) is further configured to:

check in real time, the load on the at least one node (130) when the at least one proactive action is initiated; and

halt, the proactive action once the load on the at least one node (130) is less than the second pre-defined threshold.

12. The system (120) as claimed in claim 8, wherein the load on the at least one node (130) is more than one of the first pre-defined threshold and the second pre-defined threshold during events including at least one of, unusual traffic conditions, and potential resource leakage.

13. The system (120) as claimed in claim 8, wherein the first predefined threshold and the second predefined threshold are pre-defined or dynamically set by one of the one or more processors (202).

14. The system (120) as claimed in claim 8, wherein the system (120) further comprising:

an access module (235) configured to continuously dump each type of resource usage data at pre-configured intervals into a log file.

15. A User Equipment (UE) (110), comprising:

one or more primary processors (305) communicatively coupled to one or more processors (202), the one or more primary processors (305) coupled with a memory (310), wherein said memory (310) stores instructions which when executed by the one or more primary processors (305) causes the UE (110) to:

transmit, one or more requests to the one or more processors (202) in order to avail one or more services; and

wherein the one or more processors (202) is configured to perform the steps as claimed in claim 1.

16. A non-transitory computer-readable medium having stored thereon computer-readable instructions that, when executed by a processor (205), causes the processor (205) to:

periodically monitor, load on at least one node (130) in the network (105), the load on the at least one node (130) pertains to a plurality of requests received at the at least one node (130) from one or more users associated with one or more User Equipments (UEs) (110), each request associated with availing one or more services;

raise, an alarm when the load on the at least one node (130) handling the plurality of requests is more than a first pre-defined threshold; and

initiate, at least one proactive action to balance the load, when the load on the at least one node (130) is more than a second pre-defined threshold.

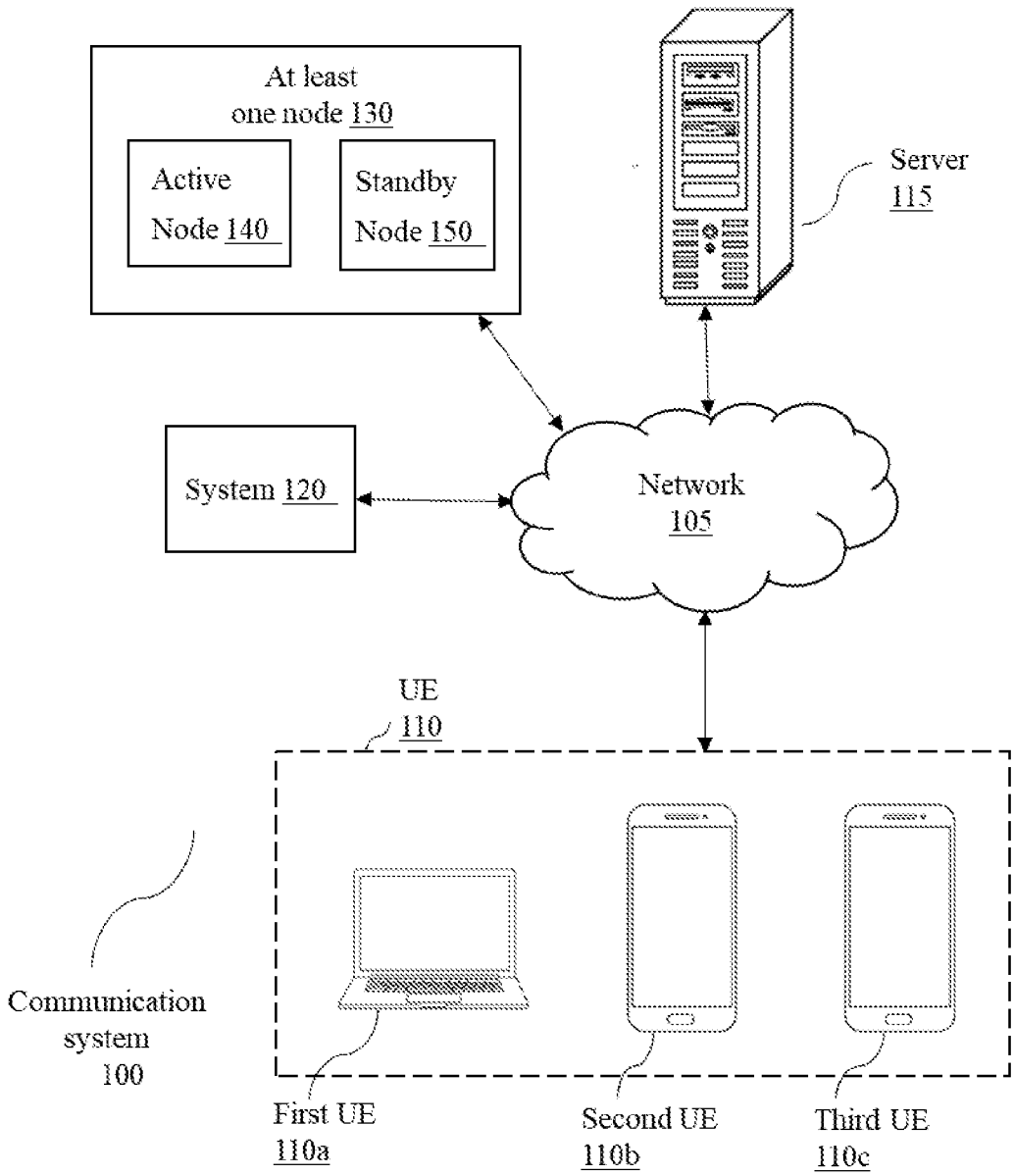
At least
one node 130
Active
Node 140
Standby
Node 150
Server
115
System 120
Network
105
UE
110
Communication
system
100
First UE
110a
Second UE
110b
Third UE
110c

FIG. 1

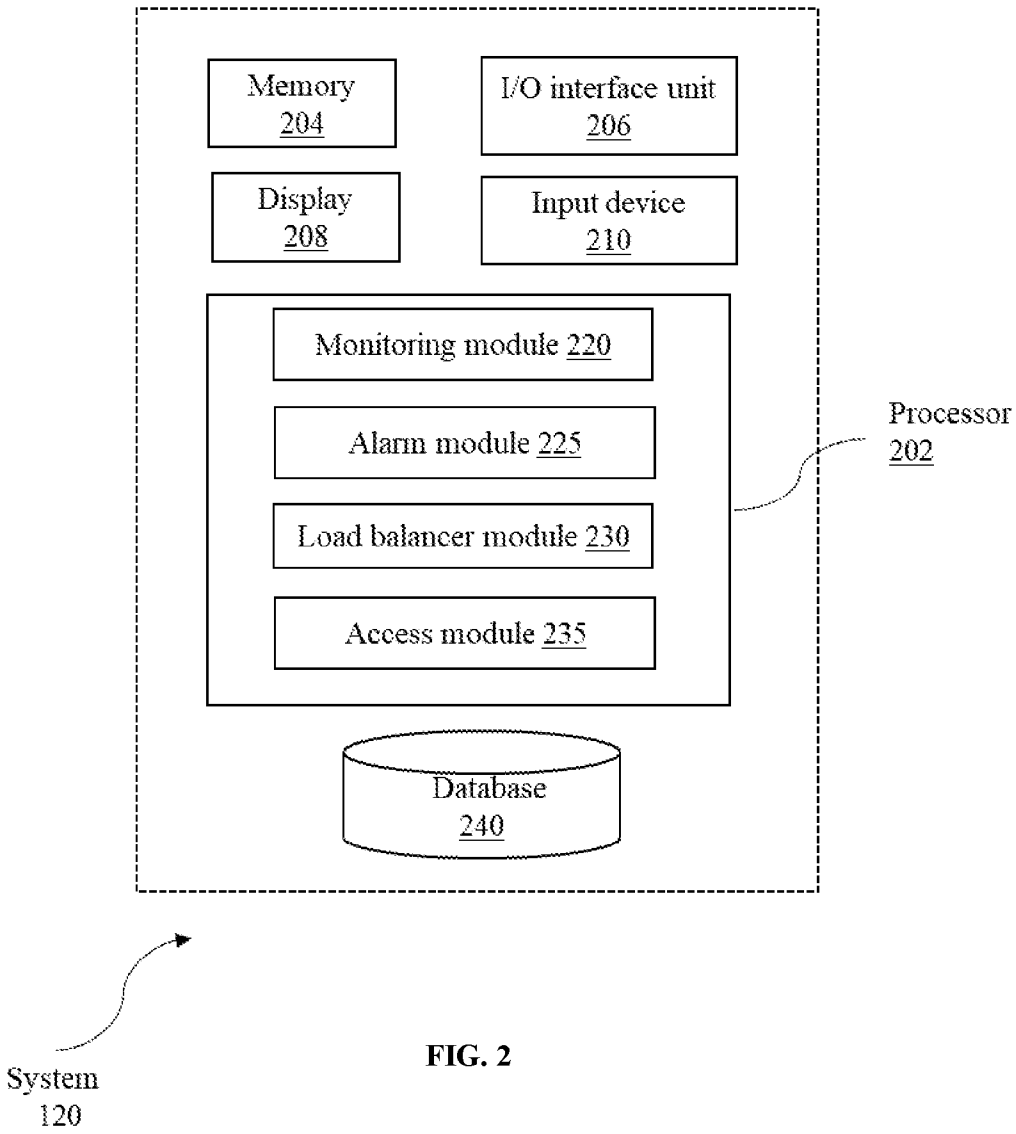
Memory
204
I/O interface unit
206
Display
208
Input device
210
Monitoring module 220
Alarm module 225
Load balancer module 230
Access module 235
Database
240
Processor
202
System
120

FIG. 2

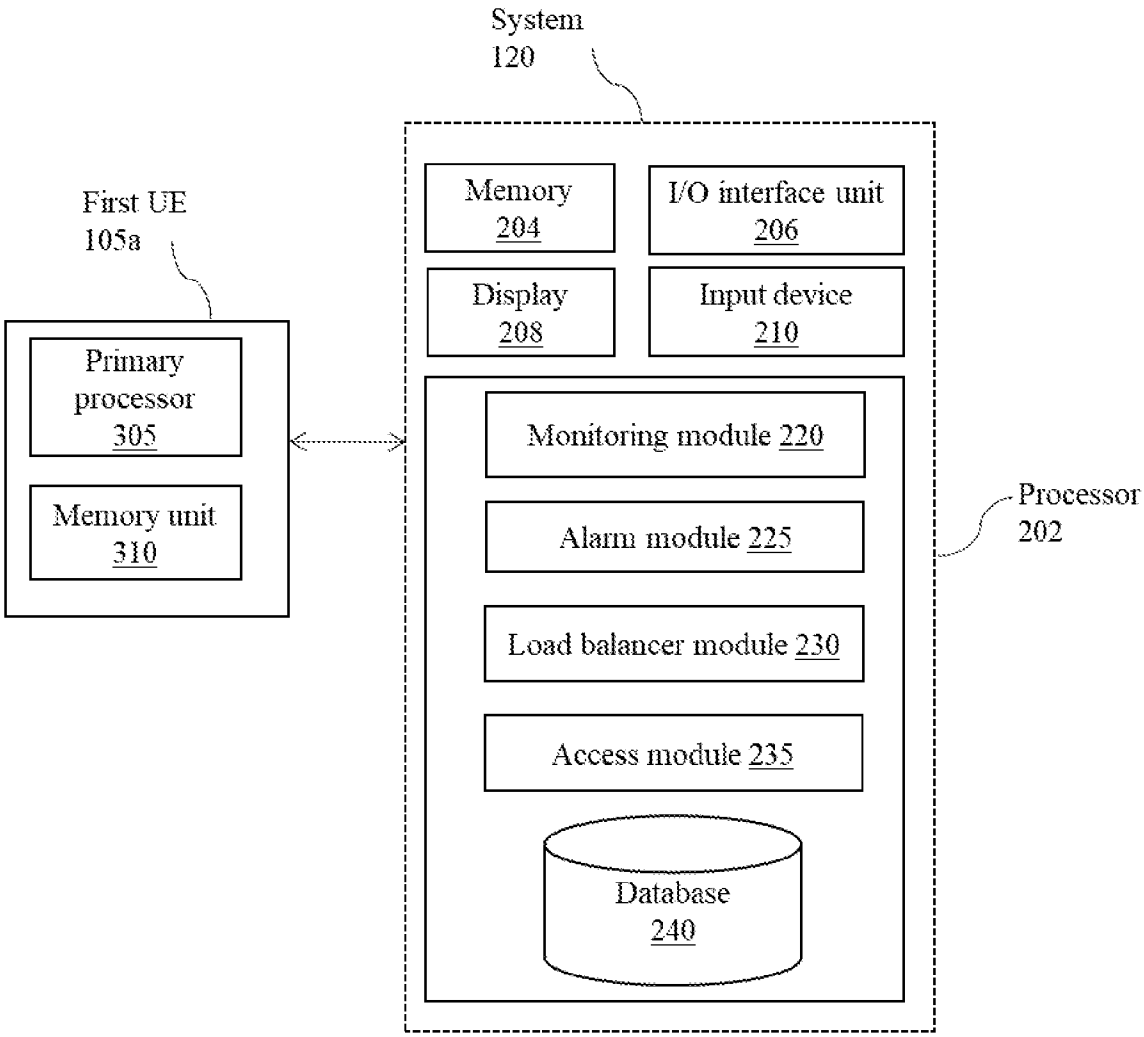
System
120
First UE
105a
Primary
processor
305
Memory unit
310
Memory
204
I/O interface unit
206
Display
208
Input device
210
Monitoring module 220
Alarm module 225
Load balancer module 230
Access module 235
Database
240
Processor
202

FIG. 3

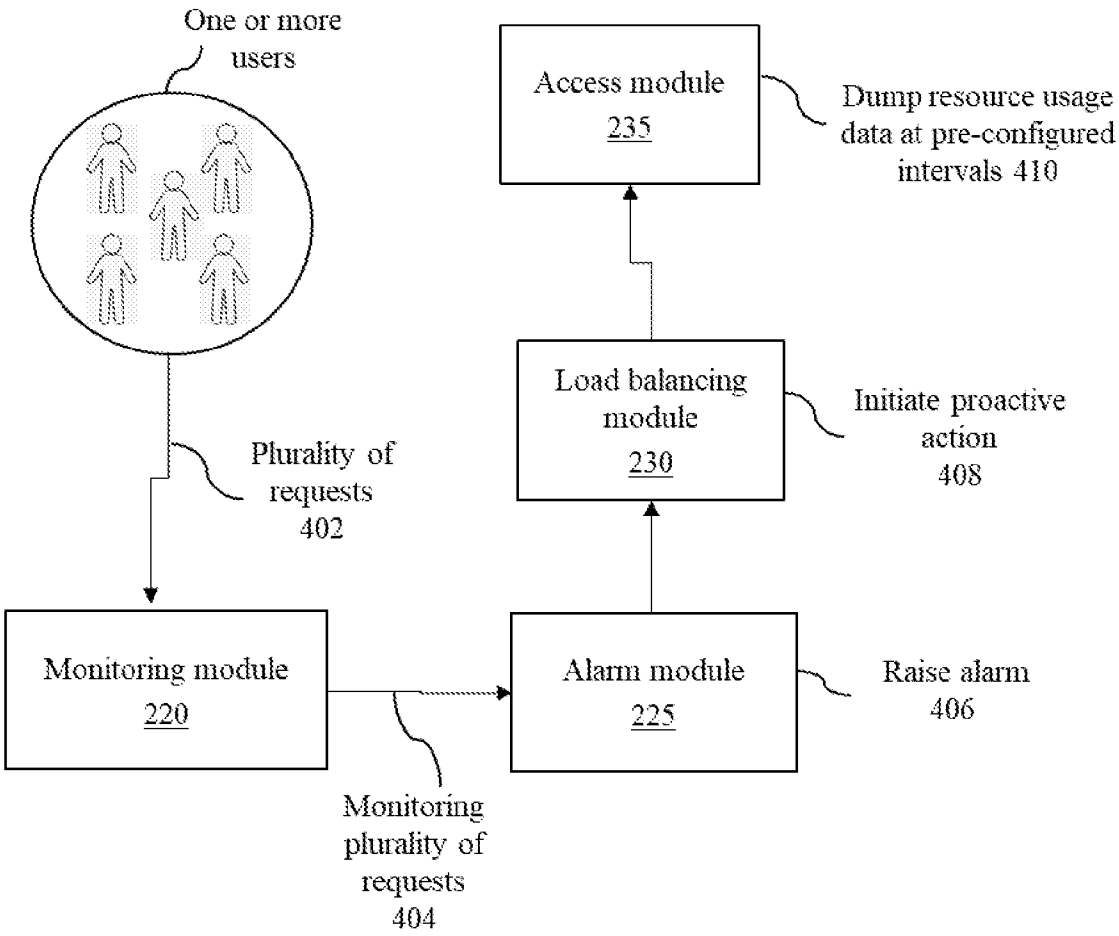
One or more users
Access module 235
Dump resource usage data at pre-configured intervals 410
Load balancing module 230
Initiate proactive action 408
Plurality of requests 402
Monitoring module 220
Alarm module 225
Raise alarm 406
Monitoring plurality of requests 404

FIG. 4

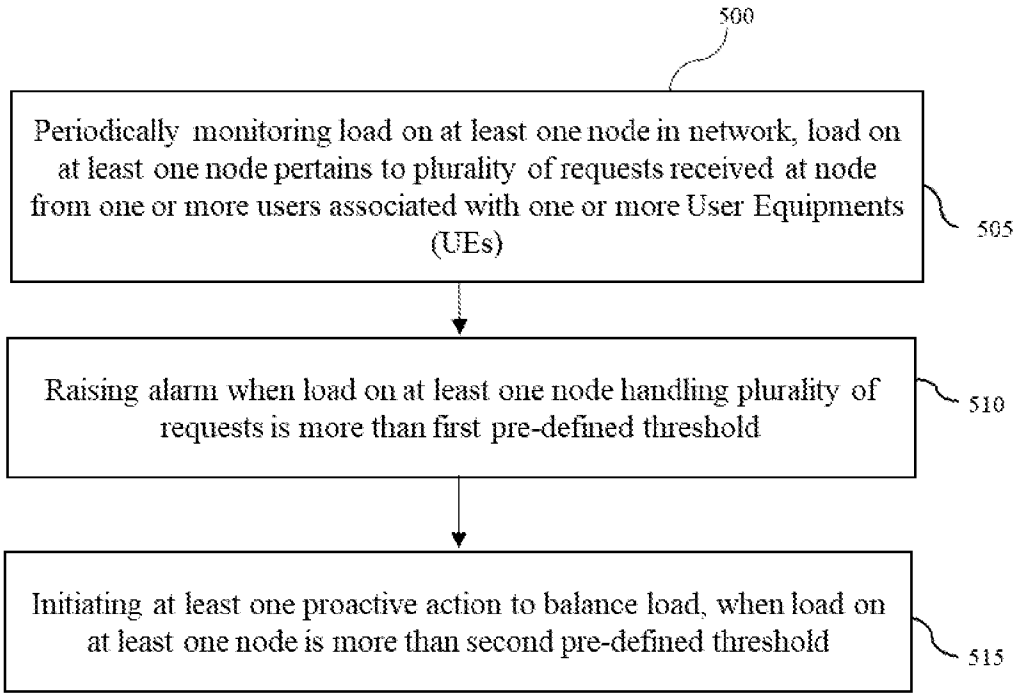
500
Periodically monitoring load on at least one node in network, load on at least one node pertains to plurality of requests received at node from one or more users associated with one or more User Equipments (UEs)
505
Raising alarm when load on at least one node handling plurality of requests is more than first pre-defined threshold
510
Initiating at least one proactive action to balance load, when load on at least one node is more than second pre-defined threshold
515

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IN2024/050931

A. CLASSIFICATION OF SUBJECT MATTER
H04L67/1008,H04L47/125,H04W24/02 Version=2024.01

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L,H04W

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

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

PatSeer, IPO Internal Database

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN105991458 B (ZTE CORP. [CN]) 17 December 2019 (17-12-2019) Google translation has been used for citation purpose. Abstract; Claim 1; Claim 4; "disclosure of Invention"	1-16
Y	CN104735724 B (ZTE CORP. [CN]) 02 July 2019 (02-07-2019) Google translation has been used for citation purpose. Abstract	1-16

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"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
28-10-2024	28-10-2024

Name and mailing address of the ISA/	Authorized officer
Indian Patent Office Plot No.32, Sector 14,Dwarka,New Delhi-110075 Facsimile No.	Amit Kumar Gupta Telephone No. +91-1125300200

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/IN2024/050931

Citation	Pub.Date	Family	Pub.Date
CN 105991458 B	17-12-2019	WO 2016123972 A1	11-08-2016
CN 104735724 B	02-07-2019	WO 2015089990 A1	25-06-2015