(19) DANMARK

Patent- og
Varemærkestyrelsen

(10) **DK/EP 3427441 T3**

(12) Oversættelse af
europæisk patentskrift

(51) Int.Cl.: ***H 04 L 12/911 (2013.01)*** *H 04 L 12/735 (2013.01)* ***H 04 L 12/703 (2013.01)*** *H 04 L 12/803 (2013.01)* ***H 04 L 12/707 (2013.01)***

(45) Oversættelsen bekendtgjort den: **2021-06-28**

(80) Dato for Den Europæiske Patentmyndigheds
bekendtgørelse om meddelelse af patentet: **2021-05-05**

(86) Europæisk ansøgning nr.: **16709054.7**

(86) Europæisk indleveringsdag: **2016-03-09**

(87) Den europæiske ansøgnings publiceringsdag: **2019-01-16**

(86) International ansøgning nr.: **EP2016055066**

(87) Internationalt publikationsnr.: **WO2017152976**

(84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(73) Patenthaver: **Telefonaktiebolaget LM Ericsson (PUBL), , 164 83 Stockholm, Sverige**

(72) Opfinder: **LINDHEIMER, Christofer, Hemmansgatan 86, 583 36 Linköping, Sverige**
VIKBERG, Jari, Svalsätersvägen 12, 153 38 Järna, Sverige

(74) Fuldmægtig i Danmark: **Zacco Denmark A/S, Arne Jacobsens Allé 15, 2300 København S, Danmark**

(54) Benævnelse: **Trafiktilgængelighed i et cellulært kommunikationsnetværk**

(56) Fremdragne publikationer:
US-A1- 2003 204 616
US-A1- 2008 259 853

DESCRIPTION

TECHNICAL FIELD

[0001] The invention relates to methods, computer programs, computer program products, network nodes and wireless devices for traffic availability in a cellular communication network.

BACKGROUND

[0002] Wireless communication in cellular networks is constantly evolving. At the moment, the communications industry puts a lot of focus on what is commonly referred to as "5G" (fifth generation) cellular networks. However, what 5G actually is, or will be is not completely clear at the moment. Part of the discussions concern what 5G should be able to do, or rather, what types of services and what performance that could be expected from 5G. This discussion connects to various use cases or usage scenarios that are envisioned. While there are still a lot of focus on quite obvious such use cases, such as web browsing, streaming services, video/ audio communication and such, there are also quite different ones. One such example is sometimes referred to as Ultra-Reliability and Ultra-Low Latency Communication, or UraLLC, for short. In this category of use cases, we find, for example, communication required to monitor and manage power grids or other critical installations, to remotely steer vehicles and machines or to perform remote surgery on patients. In these scenarios there is virtually no room for errors or other unpredictable behaviour. The connection between controller and executing entities simply has to work.

[0003] In the prior art, there is the concept of quality of service (QoS). However, QoS metrics generally refer to measures like error rates, bit rates, throughput, transmission delay, jitter etc. In summary, the current QoS relates only refer to a performance quality which is not sufficient for UraLLC.

[0004] Document D1 (US 2003/0204616 A1) discloses methods for mobile ad hoc networking. An originating mobile device sends to its neighbors a request for a communication path fulfilling certain requirements. Each receiving node propagates the request to its neighbors, until it finally reaches the requested destination. The destination then sends a reply to the originating device informing about the paths that were found. The originating device then selects to use one or more of these.

SUMMARY

[0005] It is an object to provide support to secure traffic availability in a cellular communication network according to the claims.

[0006] Link is to be construed as a communication channel between two network nodes.

[0007] Connection is to be construed as a communication path between one communication party and another, including any network nodes and links therebetween.

[0008] Generally, all terms used in the claims are to be interpreted according to their ordinary meaning in the technical field, unless explicitly defined otherwise herein. All references to "a/an/the element, apparatus, component, means, step, etc." are to be interpreted openly as referring to at least one instance of the element, apparatus, component, means, step, etc., unless explicitly stated otherwise. The steps of any method disclosed herein do not have to be performed in the exact order disclosed, unless explicitly stated.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The invention is now described, by way of example, with reference to the accompanying drawings, in which:

Fig 1 is a schematic diagram illustrating a cellular communication network where embodiments presented herein may be applied;

Fig 2 is a schematic diagram illustrating communication between a wireless device and a resource;

Fig 3 is a schematic diagram illustrating a more complex scenario of communication paths;

Fig 4 is a sequence diagram illustrating communication between nodes to secure traffic availability;

Figs 5A-D are flow charts illustrating embodiments of methods performed in any one of the network nodes of the cellular communication network of Fig 1;

Figs 6A-B are flow charts illustrating embodiments of methods performed in the wireless device of Fig 1;

Fig 7 is a schematic diagram illustrating components of any of the network nodes of Fig 1;

Fig 8 is a schematic diagram showing functional modules of the network node of Fig 7 according to one embodiment;

Fig 9 is a schematic diagram illustrating components of any of the wireless device of Fig 1;

Fig 10 is a schematic diagram showing functional modules of the wireless device of Fig 9 according to one embodiment; and

Fig 11 shows one example of a computer program product comprising computer readable

means.

DETAILED DESCRIPTION

[0010] The invention will now be described more fully hereinafter with reference to the accompanying drawings, in which certain embodiments of the invention are shown. This invention may, however, be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided by way of example so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like numbers refer to like elements throughout the description.

[0011] Using embodiments presented herein, it is possible to both represent traffic availability for a connection as well as provide the opportunity to put requirements on a connection, e.g., at connection set-up. The traffic availability expands well beyond QoS measurements by relating to how likely the connection will remain operable, giving a complete new dimension of ensuring reliability required for ultra reliable operation, such as remote surgery, remote vehicle or machine operation, etc. In addition, the different embodiments describe how availability requirements can be evaluated and calculated.

[0012] Fig 1 is a schematic diagram illustrating a cellular communication network 8 where embodiments presented herein may be applied. The cellular communication network 8 comprises a core network 3 and one or more radio access nodes 1, here e.g. in the form of radio base stations being evolved Node Bs, also known as eNode Bs or eNBs. The radio access node 1 is a network node and could also be in the form of Node Bs, BTSs (Base Transceiver Stations) and/or BSSs (Base Station Subsystems), WLAN (Wireless Local Area Network) access points, etc. The radio access node 1 provides radio connectivity over a wireless interface 4a-b to a plurality of wireless devices 2. The term wireless device is also known as mobile communication terminal, user equipment (UE), mobile terminal, user terminal, user agent, wireless terminal, machine-to-machine device etc., and can be, for example, what today are commonly known as a mobile phone, smart phone or a tablet/laptop with wireless connectivity. The term wireless is here to be construed as having the ability to perform wireless communication. More specifically, the wireless device 2 can comprise a number of wires for internal and/or external purposes.

[0013] The cellular communication network 8 may e.g. comply with any one or a combination of LTE (Long Term Evolution), W-CDMA (Wideband Code Division Multiplex), EDGE (Enhanced Data Rates for GSM (Global System for Mobile communication) Evolution), GPRS (General Packet Radio Service), CDMA2000 (Code Division Multiple Access 2000), or any other current or future wireless network, such as LTE-Advanced, as long as the principles described hereinafter are applicable. Further, the radio access part of the communication

network 8 may also include access according to other radio access technologies, e.g. any of the WLAN standards IEEE 802.11x, including physical and Media Access Control specifications. Specifically, it is currently not specified what interfaces and components that would be included in a 5G network architecture, nor how such components would be implemented. For example, functionality that usually is common to relate to, e.g., a physical node in a radio network, like mobility functions in an eNB, may very well be implemented as a virtual function on a platform or physical machine that is not part of physical equipment of the radio network. Also, in a 5G timeframe, it is not unlikely that certain portions of the complete communications network will not directly map to specific physical nodes, but rather find their realization as a logical aggregate of network functionality that utilize computing processes on various processing machines. Embodiments presented herein is to be considered applicable for any such network, irrespective of the network structure and implementation.

[0014] Over the wireless interface, uplink (UL) communication 4a occurs from the wireless device 2 to the radio access node 1 and downlink (DL) communication 4b occurs from the radio access node 1 to the wireless device 2. The quality of the wireless radio interface to each wireless device 2 can vary over time and depending on the position of the wireless device 2, due to effects such as fading, multipath propagation, interference, etc.

[0015] The radio access node 1 is also connected to the core network 3 for connectivity to central functions and a wide area network 7, such as the Internet.

[0016] The core network 3 comprises a number of network nodes used for various parts of the user plane and the control plane. The user plane is used for the transfer of user data to and from the wireless device 2. The control plane contains resources for controlling traffic in the network, allowing for a strict separation between the control plane and user plane. In Fig 1, the subset of network nodes in the core network 3 are shown: an SGW (Serving Gateway) 5, a PDN GW (Packet Data Network Gateway) 6, an MME (Mobility Management Entity) 16, a PCRF (Policy Control and Charging Rules Function) 17 and an HSS (Home Subscriber Server) 18.

[0017] The SGW 5 is a network node which acts as the anchor point for wireless device mobility, and also includes other functionalities such as temporary DL data buffering while the wireless device is being paged, packet routing and forwarding to the right eNB.

[0018] The PDN-GW 6 is the network node responsible for IP (Internet Protocol) address allocation for the wireless device 2, as well as Quality of Service (QoS) enforcement.

[0019] The MME 16 is a network node being a control node for the access network, responsible for tasks such as idle mode tracking of the wireless devices 2, paging, retransmission, bearer activation/deactivation, etc.

[0020] The PCRF 17 determines policy rules in real-time with respect to the wireless devices 2 of the system. This may e.g. include aggregating information in real-time to and from the core

network and operational support systems, etc. of the system so as to support the creation of rules and/or automatically making policy decisions for user radio terminals currently active in the system based on such rules or similar.

[0021] The HSS 17 is the network node which contains the subscription related information in order to support handling of calls and/or data sessions.

[0022] In the wide area network 7, there may be network nodes 10 a-b being intermediate network nodes, such as routers, for forwarding packets. Using the cellular communication network 8 and wide area network 7, wireless devices 2 can communicate with other resources 9 available through the wide area network. Such as resource 9 can e.g. be a server or another wireless device connected to another cellular communication network.

[0023] Fig 2 is a schematic diagram illustrating communication between a wireless device 2 and a resource 9.

[0024] The communication is here represented by a connection 13 which is a way for the network between the wireless device 2 and the resource 9 to provide an ability of communication. The connection 13 comprises all network nodes (in the control pane and/ or user plane) and links required to enable the communication. In the example of Fig 2, the connection is made up of two redundant communication paths: a first communication path 12a and a second communication path 12b. In this way, if one or more network nodes or links of the first communication path 12a were to fail, there is still a connection through the connection 13 using the second communication path 12b, albeit possibly with reduced bandwidth.

[0025] Fig 3 is a schematic diagram illustrating a more complex scenario of communication paths illustrating common nodes between redundant communication paths.

[0026] In this example, the left hand side represents one end of a connection and the right hand side represents the other end of a connection. On the left, there are three radio access nodes la-c, two SGWs 5a-b and two PDN-GWs 6a-b. Moreover, there is an Access Controller AC 15 that may control one or a set of WLAN Access Points AP's 1c, for communicating over a wireless link. On the right hand side, there are also three radio access nodes 1a'-1c', two SGWs 5a'-5b', two PDN-GWs 6a'-6b' and an AC 15'.

[0027] In the middle, between the two sides, there are a number of routers 10a-10f.

[0028] A first communication path 12a passes through the first radio access node 1a on the left, a first SGW 5a on the left, a first PDN-GW 6a on the left, the fourth router 10d, the first PDN-GW 6a' on the right, the first SGW 5a' on the right and the first radio access node 1a' on the right.

[0029] A second communication path 12a passes through the second radio access node 1b on the left, a second SGW 5b on the left, a second PDN-GW 6b on the left, the fourth router 10d,

the second PDN-GW 6b' on the right, the second SGW 5b' on the right and the second radio access node 1b' on the right.

[0030] A third communication path 12c passes through the third radio access node 1c on the left, the AC 15 on the left, the fifth router 10e, the seventh router 10g, the AC 15' on the right and the third radio access node 1c' on the right.

[0031] It can thus be seen that the first communication path 12a and the second communication path 12b have one common network node in the fourth router 10d. Hence, if the fourth router were to fail, this would affect both the first communication path 12a and the second communication path 12b negatively. Nevertheless, all other network nodes and links of the two communication paths 12a-b are separate, whereby the two paths 12a-b provide a much better traffic availability when used together. One parameter which can be used to describe this situation is a common node counter. For the communication paths 12a-c in this example, there is one common node (10d) of the connection, whereby the common node counter has the value of one.

[0032] Fig 4 is a sequence diagram illustrating communication in between nodes to secure traffic availability. The network node performing the evaluation is the evaluator 26. The requestor 25 can be another network node or a wireless device. The requestee 27 can be another network node or a wireless device.

[0033] The requestor first transmits a request 20 to the evaluator to secure traffic availability for a connection. The evaluator 26 checks internal capacity to evaluate 21 whether the evaluator 26 can internally secure requested traffic availability. The evaluator also obtains external capacity information by sending a request 20' to the next network node, the requestee 27, in the communication path of the connection. The requestee 27 provides a response 22' to the evaluator, after which the evaluator 26 can perform an external evaluation 23. Regarding the traffic availability for the link between the evaluator 26 and the requestee 27, the traffic availability of this link can be checked either by the evaluator 26 or the requestee 27, as long as the responsibility is known beforehand.

[0034] Once the evaluator 26 has evaluated both internal and external traffic availability, the evaluator 26 determines whether it can meet the traffic availability of the request 20 from the requester 25 by combining the result from the internal determination 21 and the result from the external evaluation 23, and transmits a corresponding response 22. If both of the internal and external evaluations are positive, the response 22 to the requestor 25 is positive. On the other hand, if one or both of the internal and external evaluations are negative, the response 22 to the requestor 25 is negative.

[0035] It is to be noted that the requestee 27 in turn can act as an evaluator 26. Hence, the sequence shown in Fig 4 can be employed as a recursive method, proceeding through all network nodes of the communication path(s). In this way, the availability can be secured all along the communication path(s).

[0036] Figs 5A-D are flow charts illustrating embodiments of methods performed in any one of the network nodes 1, la-c, 1a'-c', 5, 5a-b, 5a'-b', 6, 6a-b, 6a'-b', 10a-g, 15, 15', 16, 17, 18 of the cellular communication network 8 of Fig 1. First, embodiments of the method illustrated in Fig 5A will be described.

[0037] In an *obtain indicator* step 40, an indicator to secure a traffic availability for a connection is obtained. The traffic availability is related to how likely the connection will remain operable. The traffic availability thus reflects a robustness of the connection.

[0038] One example of a traffic availability parameter is an availability parameter AVA1 indicating a number of redundant communication paths between end nodes. The number of redundant communication paths is of interest from a robustness perspective, as it may indicate that if one redundant communication path fails, there is another link that can be used to continue communication. It is likely that the requirements of number of redundant communication paths may be directly related to requirements of, e.g., all the other availability parameters described below. If the other availability parameters show unreliable performance, increasing the amount of redundant communication paths may provide for an overall acceptable availability and increased reliability.

[0039] The traffic availability can comprises a plurality of availability parameters to reflect different aspects of traffic availability. For instance, the traffic availability may comprise any one or more of the following exemplifying parameters AVA2-AVA11, as well as AVA1 defined above:

AVA2: a parameter indicating maximum relative processor load of traversed network nodes. Max relative processor load of traversed nodes is an indication of how heavily loaded network nodes are. This is informational based on the assumption that the higher the load in a network node, the higher is the likelihood that a failure or even further added traffic may impact the performance related QoS and is thus informational from an overall reliability perspective.

AVA3: a parameter indicating maximum relative link load of traversed links. This is similar reason to AVA2, but instead it is now a link load that is evaluated. If a link between two nodes that is part of the connection path is heavily loaded, again, this increases the sensitivity to that further load or changes in the network may impact an ongoing connection.

AVA4: a parameter indicating operable power backup of traversed network nodes. AVA4 is a Boolean value indicating of how well a node may cope with electric power failure, if there is a back-up for traversed nodes.

AVA5: a parameter indicating a rate of packets which are passed through traversed network nodes.

AVA6: a parameter indicating a rate of packets which are passed through traversed links. AVA5 and AVA6 may be defined as the number of sent packets, divided by the total number of packets in existence in a transmission buffer.

AVA7: a parameter indicating a proportion of time that traversed network nodes are operable.

AVA8: a parameter indicating a proportion of time that traversed links are operable. AVA7 and AVA8 could also be the complement of the definitions above, i.e. a fraction of time that a network node or link is out-of-service.

AVA9: number of network node breaks per time unit (e.g. per day)

AVA10: number of node breaks per time unit (e.g. per day). AVA9 and AVA10 are thus indications of recent amount of failures.

AVA11: maximum acceptable number of common nodes (see description with reference to Fig 3 above). This parameter can be a Boolean to indicate the need of complete separation between communication paths (i.e. no common nodes) or acceptance of common nodes. Alternatively, this parameter can be an integer indicating the number of acceptable common nodes. This AVA11 is only applicable if AVA1 is greater than one.

[0040] It should be noted that AVA1-11 are exemplary availability parameters that relate to availability and thus the reliability of a connection between two entities. Other availability parameters may be relevant, e.g., in particular in network architectures where, e.g., features and functionality are all software defined and realized using processing resources in the cloud. In such situations, a relevant availability parameters may go more towards how many control functions that share the same control interface towards a radio node or similar. Further, if a network node in a sense is in fact made up of a number of processing units that may even be distributed geographically, or be cloud-based etc., a virtual node availability parameters may be constructed based on a worst-case value of components necessary to make up the node. Thus, an availability parameters may propagate down to processing/CPU/Memory level and be aggregated upwards with the weakest-link principle that it is always the worst availability value that is the one being propagated through the system.

[0041] The indicator to secure traffic availability can be obtained using one or more of the following.

[0042] The indicator can be received explicitly in a request (see 20 of Fig 4) comprising the indicator.

[0043] Alternatively or additionally, the indicator can be based on a wireless device of one end of the connection. For instance, the wireless device can be associated with a predetermined traffic availability using its subscription. For instance, police officers can be configured to always receive a traffic availability of a certain level based on their subscription data.

[0044] Alternatively or additionally, the indicator can be based on a network slice associated with the connection. Each network slice can have its own set of (physical and/or virtual) resources in the core network, while radio access network resources can be shared between

network slices. For instance, a specific network slice can be established for remote surgery, and when a session is set up using the remote surgery network slice, a very high traffic availability is determined, compared e.g. to a media streaming network slice.

[0045] Alternatively or additionally, the indicator can be based on a user service identifier for the wireless device. For instance, deep packet inspection can be used to identify the user service and determine the traffic availability based on the user service.

[0046] In a *determine resources* step 42, it is determined whether the network node can secure resources to support the traffic availability.

[0047] In a conditional *resources* step 43, the method proceeds conditionally based on the determination in step 42. If the network node can secure resources, the method proceeds to an *allocate* step 44. Otherwise, the method proceeds to a *transmit negative response* step 45.

[0048] In the *allocate* step 44, resources are allocated for the connection when resources can support the traffic availability.

[0049] In a *transmit positive response* step 46, a positive response is transmitted.

[0050] In the *transmit negative response* step 45, a negative response is transmitted.

[0051] Looking now to Fig 5B, only new or modified steps compared to Fig 5A will be described.

[0052] In an optional *determine lower traffic availability* step 48, a lower traffic availability that can be secured is determined. The lower traffic availability is lower than the traffic availability of the indicator, since this step only is performed when there are not sufficient resources to secure the requested traffic availability.

[0053] In this case, the *transmit negative response* step 45, optionally comprises transmitting a negative response comprising the lower traffic availability. In this way, the requester can request a new traffic availability which the network node can secure. In other words, the method is then repeated for a new request comprising an indicator to secure (at most) the lower traffic availability.

[0054] In an optional conditional *negative event* step 47, it is determined whether the network node detects an event which negatively affects traffic availability of the connection. This may be the situation for example, if a connection including one or more redundant communication paths experience that one link fails, or that one node used by one of the communication paths is reaching a high-load situation (higher than what was originally negotiated/accepted) and that congestion actions may start occurring. In such situations, connection entities need to be notified of such a situation, even though the communication is still reaching the end points, since the reliability level is at risk.

[0055] If a negative event is detected, the method proceeds to a *transmit warning signal* step 49. Otherwise (when polling is used) the step is repeated, optionally after a delay (not shown).

[0056] In the *transmit warning signal* step 49, a warning signal is transmitted. The warning signal can be issued on the protocol level where it is detected, e.g., IP level for switching nodes, and thus propagate to the end-points using similar translation engines (see below) used for setting up the connection and negotiating the current availability level. It would then further be possible to re-negotiate the availability, or to terminate the connection, or to attempt setup of additional redundant communication paths to meet the requirements of the connection.

[0057] Looking now to Fig 5C, optional substeps of the *determine resources* step 42 are described.

[0058] In an optional *determine internal resources* step 42a, the network node determined whether it can secure internal resources to support the traffic availability. This corresponds to step 21 of Fig 4.

[0059] In an optional *determine external resources* step 42b, the network node determines external capability by determining whether a link connected to the network node and/or another network node can secure external resources to support the traffic availability.

[0060] When both internal resources and external resources can be secured, the network node thus determines that resources can be secured.

[0061] Looking now to Fig 5D, optional substeps of the *determine external resources* step 42b are described.

[0062] In an optional *transmit request* step 42c, a request is transmitted to an external resource to secure the traffic availability for the connection. This corresponds to message 20' of Fig 4.

[0063] In an optional *receive response* step 42d, a response is received from the external resource indicating whether the external resource can support the traffic availability. This corresponds to message 22' of Fig 4.

[0064] The additions suggested above that relate to redundancy requirements for established connections can be illustrated using the QoS concept for EPS (Evolved Packet System) bearers. To get an end-to-end QoS treatment also including availability, similar QoS extensions are added to IP schemes that relate to QoS, e.g. DiffServ, RSVP (Resource Reservation Protocol), IntServ, RSVP-TE (RSVP Traffic Engineering), or NSLP (NSIS (Next Steps in Signaling) Signaling Layer Protocol). If these are not done in a way that is directly translatable (1-to-1-mapping) a translation function of such QoS reliability primitives is provided, in a similar

way as is done for, e.g., VoLTE (Voice over LTE) with QCI (QoS Class Indicator) to Diff-Serv mapping.

[0065] The interpretation and mapping of the traffic availability when related to a network involving EPC could be implemented as part of the PCRF functionality and mapping between different availability representations in various protocols can be specified in a similar manner as for, e.g. VoLTE, between IP flows over Internet and EPS bearers.

[0066] An alternative way of controlling the availability aspect of a connection if just considering the radio network part and part of the core network part after, e.g., the PDN gateway, would be to introduce QoS-Availability information over an Rx interface and insert policies for certain connections that govern what availability level that is acceptable over the part of the end-to-end connection that relates to the mobile/radio network. Such availability levels, or criteria may thus also be governed from an operator or application/ service provider perspective, through applying policies in the PCRF, e.g., for specific users, for specific services or any combination thereof. This Rx interface may also be used to impact the QoS mapping between, e.g., QoS primitives of EPS bearers with QoS primitives and representations on IP layer.

[0067] With signaling over the Rx interface to PCRF, for example, a fire alarm company, connecting their alarms through a cellular operator, may want all the alarms to have a quality-of-service level with respect to availability, such that the fire alarms are always served over at least two redundant communication paths (see above as AVA1). As explained above for step 40, certain subscriptions (i.e., the subscription to serve a number of fire alarms) may also have certain availability levels inherent and then information may come, not over Rx but instead be part of stored subscription data in HSS.

[0068] These types of policies may be communicated from an operator over the Rx interface, and it may be associated with a certain set of fire alarms only (e.g., a certain subscription). This Rx-driven request for availability means that any application function, or for that matter a detection function that has detected a certain traffic type, can send a request for a certain availability level to the PCRF. The PCRF would then propagate this further to PDN-GW, enforcing this functionality. These availability rules may thus be specified outside of the communications network, rather in the application layer.

[0069] The addition of a new availability parameter or a composite of several availability parameters can be done in a number of different ways. The QoS framework can be extended, or coding of an already existing field can be specified to point to another QoS information element, e.g., availability. In one example the availability related information, i.e. the separate availability parameters or a composite of several availability parameters is signalled from the core network to the radio access network when a radio bearer is to be established. In another example, the signaling of availability requirements is based on a specific set of QCI and in this case the radio access network is configured with the separate availability parameters or a composite of several availability parameters for example based on the QCI (that is signalled).

[0070] Now, various aspects of calculating aggregated availability will be described. At connection establishment, it is feasible to calculate an aggregated value of availability, including all components that are part of an end-to-end connection, or between a sub-part of a connection path. A value of availability can be defined by the weakest link in some sense and to represent an end-to-end availability value, a function is needed to aggregate availability parameters of different network portions.

[0071] As an example, consider a wireless device-to-wireless device communication case between two cities, where the initiating wireless device belongs to PLMN (Public Land Mobile Network) 1 and the receiving wireless device belongs to PLMN 2. In calculating an aggregate value of availability, thus, at least three different parts need to be considered.

1. A) PLMN1
2. B) Internet route between PDN GWs for PLMN 1 and PLMN2
3. C) PLMN2

[0072] For the examples availability parameters above, this could be done in the following manner:

AVA1 - number of redundant communication paths: an availability parameter would be the number of paths that are established between the peers. For example, using MPTCP (Multipath Transport Control Protocol) and both PLMN1 and WLAN, it would be two communication paths. If only a single PLMN1-PLMN2 link would be established, it would be one communication path.

AVA2 - max relative processor load of traversed nodes would be calculated as the reported or measured relative load of the most loaded traversed node.

AVA3 - max relative link load of traversed links would be calculated as the reported or measured relative load of the most loaded traversed link (also including radio links).

AVA4 - operable power backup would be a single value being true if all involved traversed network nodes report that back-up power is available and not used, otherwise it would be false.

AVA5/ 6 - would be the minimum value of the reliability rate of any of the traversed nodes/ links.

AVA7/ 8 would be the maximum value of out of service of any of the traversed nodes/ links.

AVA 9/10 would be the maximum value of link/ node breaks of any of the traversed nodes/ links.

[0073] The aggregate view could be gathered in connection to the enforcement function in the PDN Gateway.

[0074] If there are back-up connections, i.e., if the number of redundant communication paths are greater than one, one principle could be to always calculate and report the most reliable availability parameter for one and the same link, but never mix the values and pick an aggregated AVA4 value from, e.g., path 1 and an aggregated AVA3-value from path 2.

[0075] As has been described above, a number of different availability parameters are provided to include in QoS schemes, both in mobile network, i.e., for EPS or similar bearers, as well as for IP level communications, and translation functions therebetween have been mentioned. Further, with availability being a complex area and increasingly difficult to parameterize in scenarios when some of the network functionality rely on processing possibilities implemented on cloud level, an aggregated availability parameter may become inevitable and this aggregated availability parameter maybe an aggregate of., e.g., availability parameters similar to AVA1-11 as described above and it may further indicate an availability profile with AVA1-11 as components in such a profile. For example, an availability profile 1 can imply specific values with respect to AVA1, AVA2, etc. On the other hand, an availability profile 2 can imply at least one different specific value with respect to AVA1, AVA2, etc. compared to availability profile 1. At bearer establishment and for enforcement and alarm functionality, it is thus the availability profile that is requested and enforced.

[0076] Figs 6A-B are flow charts illustrating embodiments of methods performed in the wireless device 2 of Fig 1. First, embodiments of the method illustrated in Fig 6A will be described.

[0077] In a *transmit request* step 50, the wireless device transmits a request to a network node. The request comprises an indicator to secure a traffic availability for a connection. As explained above, the traffic availability is related to how likely the connection will remain operable.

[0078] In a *receive positive response* step 52, a positive response is received, indicating that the network node can support the traffic availability.

[0079] In a *receive warning* step 54, a warning signal is received when the network node detects an event which negatively affects traffic availability of the connection.

[0080] Looking now to Fig 6B, only new or modified steps compared to Fig 6A will be described.

[0081] In an optional *present warning* step 56, the wireless device presents a user warning indicating reduced traffic availability on a user output device (169 of Fig 9) of the wireless device.

[0082] In an optional *renegotiate* step 58, a new traffic availability is renegotiated with the network node.

[0083] Fig 7 is a schematic diagram illustrating components of any of the network nodes 1, la-c, 1a'-c', 5, 5a-b, 5a'-b', 6, 6a-b, 6a'-b', 10a-g, 15, 15', 16, 17, 18 of Fig 1. A processor 60 is provided using any combination of one or more of a suitable central processing unit (CPU), multiprocessor, microcontroller, digital signal processor (DSP), application specific integrated circuit etc., capable of executing software instructions 67 stored in a memory 64, which can thus be a computer program product. The processor 60 can be configured to execute the method described with reference to Figs 5A-D above.

[0084] The memory 64 can be any combination of read and write memory (RAM) and read only memory (ROM). The memory 64 also comprises persistent storage, which, for example, can be any single one or combination of magnetic memory, optical memory, solid state memory or even remotely mounted memory.

[0085] A data memory 66 is also provided for reading and/or storing data during execution of software instructions in the processor 60. The data memory 66 can be any combination of read and write memory (RAM) and read only memory (ROM).

[0086] The network node further comprises an I/O interface 62 for communicating with other external entities. Optionally, the I/O interface 62 also includes a user interface.

[0087] Other components of the network node are omitted in order not to obscure the concepts presented herein.

[0088] Fig 8 is a schematic diagram showing functional modules of the network node of Fig 7 according to one embodiment. The modules are implemented using software instructions such as a computer program executing in the network node. Alternatively or additionally, the modules are implemented using hardware, such as any one or more of an ASIC (Application Specific Integrated Circuit), an FPGA (Field Programmable Gate Array), or discrete logical circuits. The modules correspond to the steps in the methods illustrated in Figs 5A to 5D.

[0089] An obtainer 70 corresponds to step 40. A determiner 72 corresponds to steps 42, 42a, 42b, 43, 47, and 48. An allocator 74 corresponds to step 44. A transmitter 75 corresponds to steps 45, 46, 49, and 42c. A receiver 76 corresponds to step 42d.

[0090] Fig 9 is a schematic diagram illustrating components of any of the wireless device 2 of Fig 1. A processor 160 is provided using any combination of one or more of a suitable central processing unit (CPU), multiprocessor, microcontroller, digital signal processor (DSP), application specific integrated circuit etc., capable of executing software instructions 167 stored in a memory 164, which can thus be a computer program product. The processor 160 can be configured to execute the method described with reference to Figs 6A-B above.

[0091] The memory 164 can be any combination of read and write memory (RAM) and read only memory (ROM). The memory 164 also comprises persistent storage, which, for example, can be any single one or combination of magnetic memory, optical memory, solid state

memory or even remotely mounted memory.

[0092] A data memory 166 is also provided for reading and/or storing data during execution of software instructions in the processor 160. The data memory 166 can be any combination of read and write memory (RAM) and read only memory (ROM).

[0093] The wireless device further comprises an I/O interface 162 for communicating with other external entities. Also, the wireless device comprises a user interface 169 to receive input from the user and provide output to the user. For instance, the user interface 169 may comprise any one or more of the following: a display, a touch sensitive display, a speaker, a microphone, physical keys, a microphone, etc. The user interface 169 can be used to present warnings with regard to availability due to negative events.

[0094] Other components of the wireless device are omitted in order not to obscure the concepts presented herein.

[0095] Fig 10 is a schematic diagram showing functional modules of the wireless device 2 of Fig 9 according to one embodiment. The modules are implemented using software instructions such as a computer program executing in the wireless device 2. Alternatively or additionally, the modules are implemented using hardware, such as any one or more of an ASIC (Application Specific Integrated Circuit), an FPGA (Field Programmable Gate Array), or discrete logical circuits. The modules correspond to the steps in the methods illustrated in Figs 6A and 6B.

[0096] A transmitter 170 corresponds to step 50. A receiver 172 corresponds to steps 52 and 54. Apresenter 174 corresponds to step 56. Arenegotiator 176 corresponds to step 58.

[0097] Fig 11 shows one example of a computer program product comprising computer readable means. On this computer readable means a computer program 91 can be stored, which computer program can cause a processor to execute a method according to embodiments described herein. In this example, the computer program product is an optical disc, such as a CD (compact disc) or a DVD (digital versatile disc) or a Blu-Ray disc. As explained above, the computer program product could also be embodied in a memory of a device, such as the computer program product 64 of Fig 7 or the computer program product 164 of Fig 9. While the computer program 91 is here schematically shown as a track on the depicted optical disk, the computer program can be stored in any way which is suitable for the computer program product, such as a removable solid state memory, e.g. a Universal Serial Bus (USB) drive.

[0098] The invention has mainly been described above with reference to a few embodiments. However, as is readily appreciated by a person skilled in the art, other embodiments than the ones disclosed above are equally possible within the scope of the invention, as defined by the appended patent claims.

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US20030204616A1 [0004]

Patentkrav

1. Fremgangsmåde udført i en netværksknude (1, 1a-c, 1a'-c', 5, 5a-b, 5a'-b', 6, 6a-b, 6a'-b', 10a-g, 15, 15', 16, 17, 18) af et cellulært kommunikationsnetværk (8), hvor fremgangsmåden omfatter følgende trin:
at opnå (40) en indikator for at sikre en trafiktilgængelighed for en forbindelse, hvor trafiktilgængeligheden er relateret til, hvor sandsynligt det er, at forbindelsen vil forblive anvendelig;
at bestemme (42), om netværksknuden kan sikre ressourcer til at understøtte trafiktilgængeligheden;
at allokere (44) ressourcer til forbindelsen, når ressourcer kan understøtte trafiktilgængeligheden; og
at sende (46) et positivt svar, når netværksknuden kan sikre ressourcer til at understøtte trafiktilgængeligheden;
kendetegnet ved, at trafiktilgængeligheden omfatter et tilgængelighedsparameter, der angiver et antal af redundante kommunikationsveje, der er større end en, og hvor trafiktilgængeligheden omfatter en indikator for maksimalt acceptabelt antal af fælles knuder, der er fælles for mindst to af de redundante kommunikationsveje.

2. Fremgangsmåde ifølge krav 1, hvor trinnet med at opnå (40) en indikator omfatter at bestemme indikatoren baseret på en netværk-slice associeret med forbindelsen.

3. Fremgangsmåde ifølge krav 1, hvor trinnet med at opnå (40) en indikator omfatter at bestemme indikatoren baseret på en brugerserviceidentifikator til den trådløse indretning.

4. Fremgangsmåde ifølge et af de foregående krav, yderligere omfattende trinnet:
at sende (49) et advarselssignal, når netværksknuden detekterer en begivenhed, der påvirker trafiktilgængeligheden af forbindelsen negativt.

5. Netværksknude (1, 1a-c, 1a'-c', 5, 5a-b, 5a'-b', 6, 6a-b, 6a'-b', 10a-g, 15, 15', 16, 17, 18) omfattende:
midler til at opnå en indikator for at sikre en trafiktilgængelighed for en forbindelse, hvor trafiktilgængeligheden er relateret til, hvor sandsynligt det er, at forbindelsen vil forblive anvendelig;
midler til at bestemme, om netværksknuden kan sikre ressourcer til at understøtte trafiktilgængeligheden;
midler til at allokere ressourcer til forbindelsen, når ressourcer kan understøtte trafiktilgængeligheden; og
midler til at sende et positivt svar, når netværksknuden kan sikre ressourcer til at understøtte trafiktilgængeligheden;
kendetegnet ved, at trafiktilgængeligheden omfatter et tilgængelighedsparameter, der angiver et antal af redundante kommunikationsveje, der er større end en, og hvor trafiktilgængeligheden omfatter en indikator for maksimalt acceptabelt antal af fælles knuder, der er fælles for mindst to af de redundante kommunikationsveje.

6. Netværksknude ifølge krav 5, hvor midlerne til at opnå en indikator omfatter midler til at bestemme indikatoren baseret på en netværk-slice associeret med forbindelsen.

7. Netværksknude ifølge krav 5, hvor midlerne til at opnå en indikator omfatter midler til at bestemme indikatoren baseret på en brugerserviceidentifikator til den trådløse indretning.

8. Netværksknude ifølge et af kravene 5 til 7, yderligere omfattende midler til at sende et advarselssignal, når netværksknuden detekterer en begivenhed, der påvirker trafiktilgængeligheden af forbindelsen negativt.

9. Computerprogram (66, 91), hvor computerprogrammet omfatter en computerprogramkode, som, når den køres på en netværksknude (1, 1a-c, 1a'-c', 5, 5a-b, 5a'-b', 6, 6a-b, 6a'-b', 10a-g, 15, 15', 16, 17, 18) bevirker, at netværksknuden:

opnår en indikator for at sikre en trafiktilgængelighed for en forbindelse, hvor trafiktilgængeligheden er relateret til, hvor sandsynligt det er, at forbindelsen vil forblive anvendelig;
bestemmer, om netværksknuden kan sikre ressourcer til at understøtte trafiktilgængeligheden;
allokerer ressourcer til forbindelsen, når ressourcer kan understøtte trafiktilgængeligheden; og
sender et positivt svar, når netværksknuden kan sikre ressourcer til at understøtte trafiktilgængeligheden;
kendetegnet ved, at trafiktilgængeligheden omfatter et tilgængelighedsparameter, der angiver et antal af redundante kommunikationsveje, der er større end en, og hvor trafiktilgængeligheden omfatter en indikator for maksimalt acceptabelt antal af fælles knuder, der er fælles for mindst to af de redundante kommunikationsveje.

10. Fremgangsmåde udført i en trådløs indretning (2) i et cellulært kommunikationsnetværk (8), hvor fremgangsmåden omfatter følgende trin:
at sende (50) en anmodning til en netværksknude (1, 1a-c, 1a'-c', 16), hvor anmodningen omfatter en indikator for at sikre en trafiktilgængelighed for en forbindelse, hvor trafiktilgængeligheden er relateret til, hvor sandsynligt det er, at forbindelsen vil forblive anvendelig;
at modtage (52) et positivt svar, når netværksknuden kan understøtte trafiktilgængeligheden; og
at modtage (54) et advarselssignal, når netværksknuden detekterer en begivenhed, der påvirker trafiktilgængeligheden af forbindelsen negativt.
kendetegnet ved, at trafiktilgængeligheden omfatter et tilgængelighedsparameter, der angiver et antal af redundante kommunikationsveje, der er større end en, og hvor trafiktilgængeligheden omfatter en indikator for maksimalt acceptabelt antal af fælles knuder, der er fælles for mindst to af de redundante kommunikationsveje.

11. Fremgangsmåde ifølge krav 10, yderligere omfattende trinnet:
at præsentere (56) en brugeradvarsel, der angiver reduceret trafiktilgængelighed på en bruger-outputindretning af den trådløse indretning.

12. Fremgangsmåde ifølge krav 10 eller 11, yderligere omfattende trinnet:
at genforhandle (58) en ny trafiktilgængelighed med netværksknuden.

13. Trådløs indretning (2) omfattende:
midler til at sende en anmodning til en netværksknude (1, 1a-c, 1a'-c', 16), hvor anmodningen omfatter en indikator for at sikre en trafiktilgængelighed for en forbindelse, hvor trafiktilgængeligheden er relateret til, hvor sandsynligt det er, at forbindelsen vil forblive anvendelig;
midler til at modtage et positivt svar, når netværksknuden kan understøtte trafiktilgængeligheden; og
midler til at modtage et advarselssignal, når netværksknuden detekterer en begivenhed, der påvirker trafiktilgængeligheden af forbindelsen negativt.
kendetegnet ved, at trafiktilgængeligheden omfatter et tilgængelighedsparameter, der angiver et antal af redundante kommunikationsveje, der er større end en, og hvor trafiktilgængeligheden omfatter en indikator for maksimalt acceptabelt antal af fælles knuder, der er fælles for mindst to af de redundante kommunikationsveje.

14. Trådløs indretning (2) ifølge krav 13, yderligere omfattende midler til at præsentere en brugeradvarsel, der angiver reduceret trafiktilgængelighed på en bruger-outputindretning af den trådløse indretning.

15. Trådløs indretning (2) ifølge krav 13 eller 14, yderligere omfattende midler til at genforhandle en ny trafiktilgængelighed med netværksknuden.

16. Computerprogram (66, 91), hvor computerprogrammet omfatter en computerkode, som, når den køres på en trådløs indretning (2), bevirker, at den trådløse indretning (2):

sender en anmodning til en netværksknude (1, 1a-c, 1a'-c', 16), hvor anmodningen omfatter en indikator for at sikre en trafiktilgængelighed for en forbindelse, hvor trafiktilgængeligheden er relateret til, hvor sandsynligt det er, at forbindelsen vil forblive anvendelig;

modtager et positivt svar, når netværksknuden kan understøtte trafiktilgængeligheden; og

modtager et advarselssignal, når netværksknuden detekterer en begivenhed, der påvirker trafiktilgængeligheden af forbindelsen negativt.

kendetegnet ved, at trafiktilgængeligheden omfatter et tilgængelighedsparameter, der angiver et antal af redundante kommunikationsveje, der er større end en, og hvor trafiktilgængeligheden omfatter en indikator for maksimalt acceptabelt antal af fælles knuder, der er fælles for mindst to af de redundante kommunikationsveje.

DRAWINGS

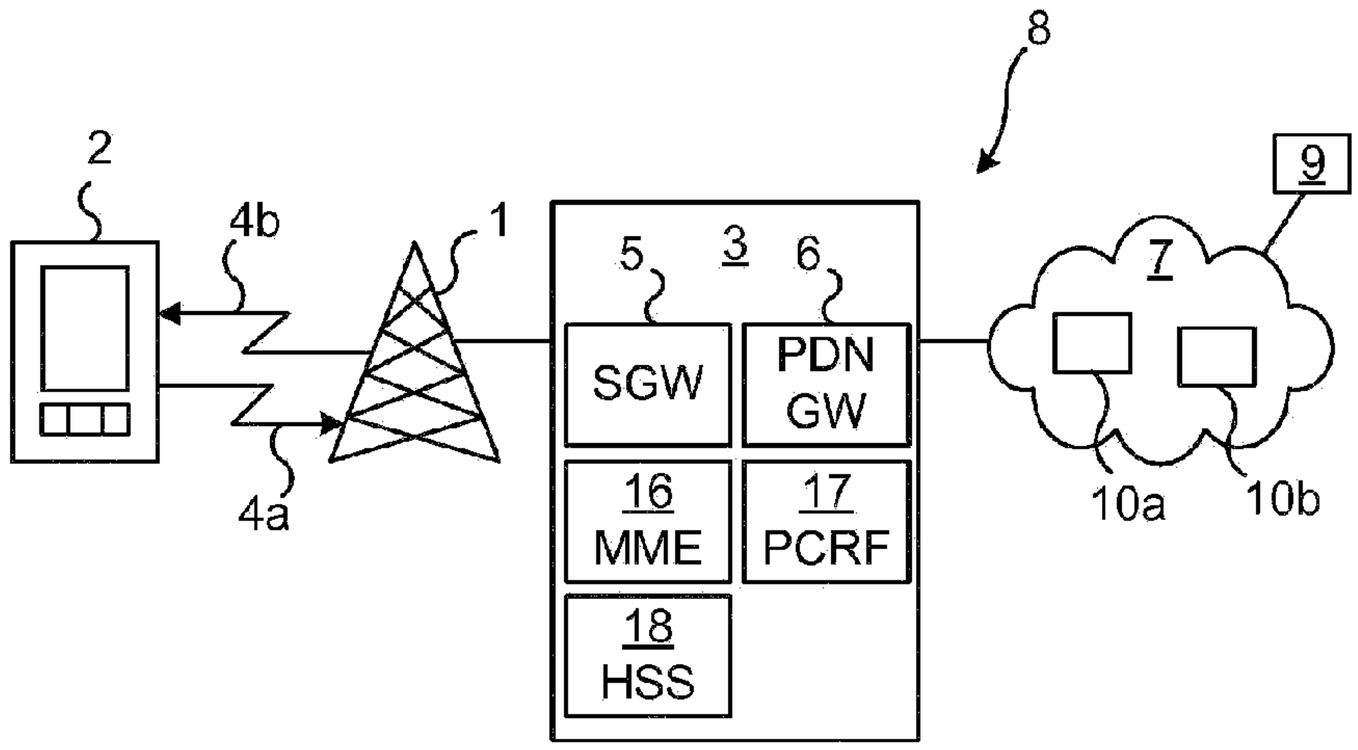

Fig. 1

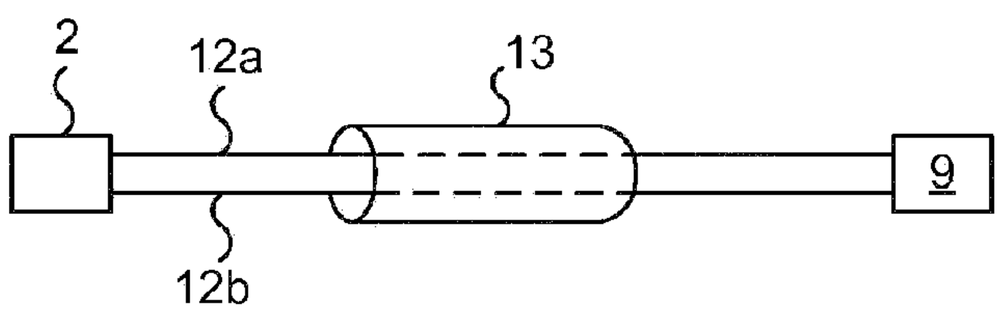

Fig. 2

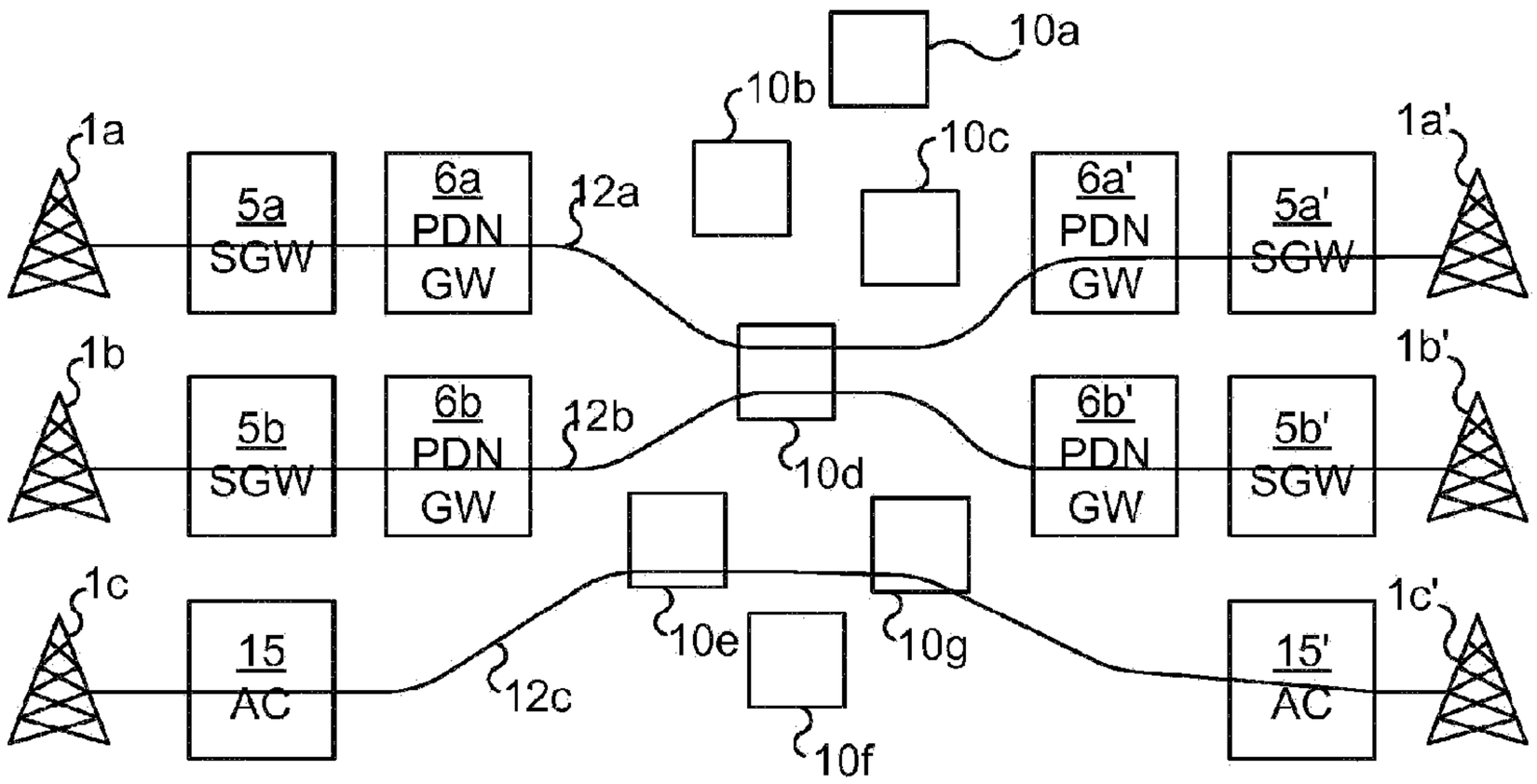

Fig. 3

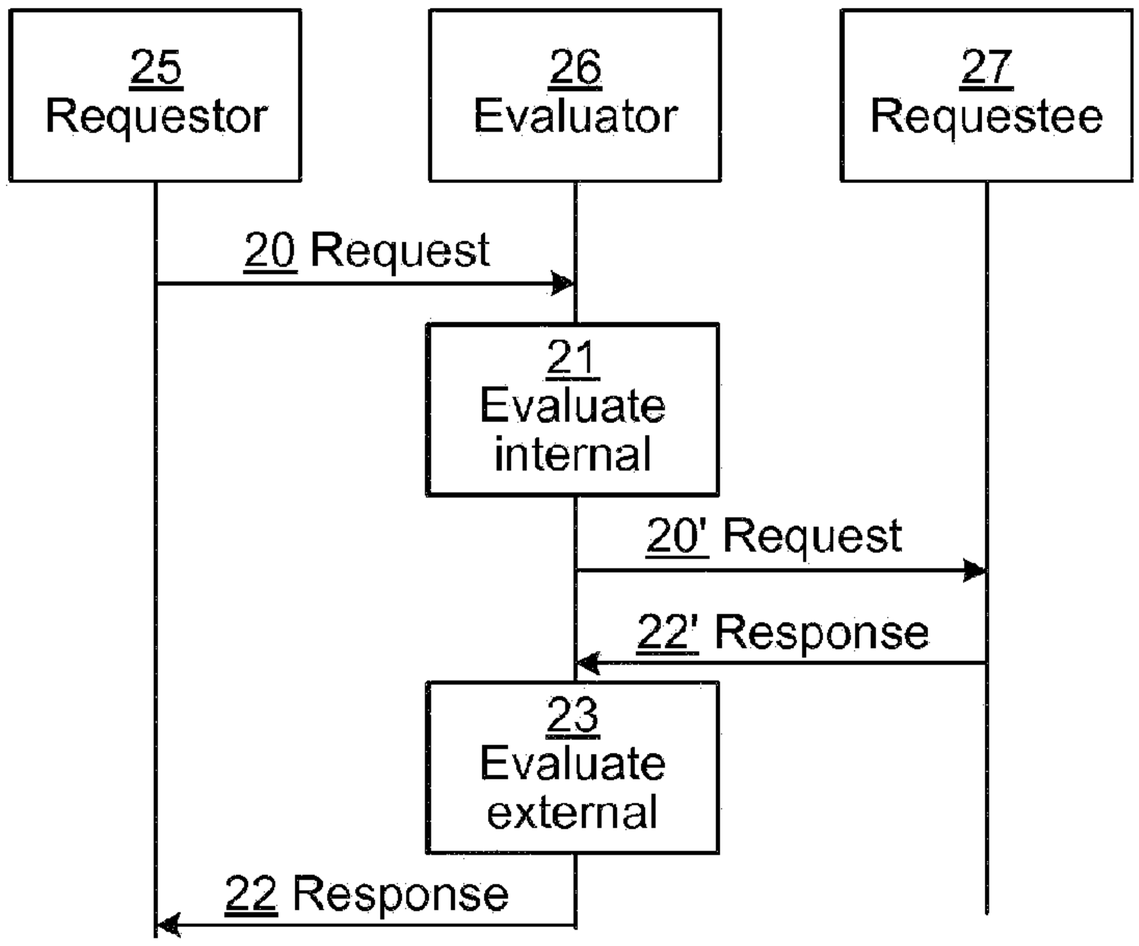
25
Requestor
26
Evaluator
27
Requestee
20 Request
21
Evaluate
internal
20' Request
22' Response
23
Evaluate
external
22 Response

Fig. 4

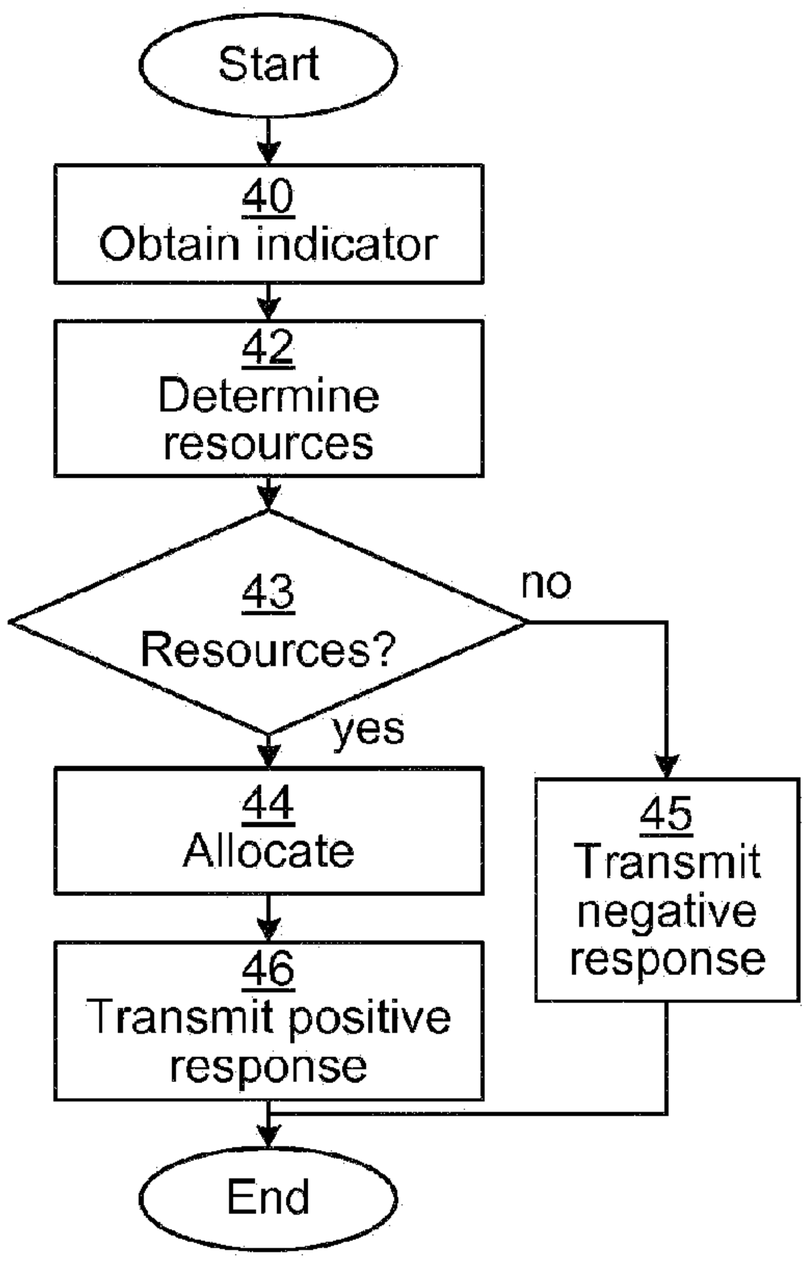
Start
40
Obtain indicator
42
Determine
resources
43
Resources?
no
yes
44
Allocate
45
Transmit
negative
response
46
Transmit positive
response
End

Fig. 5A

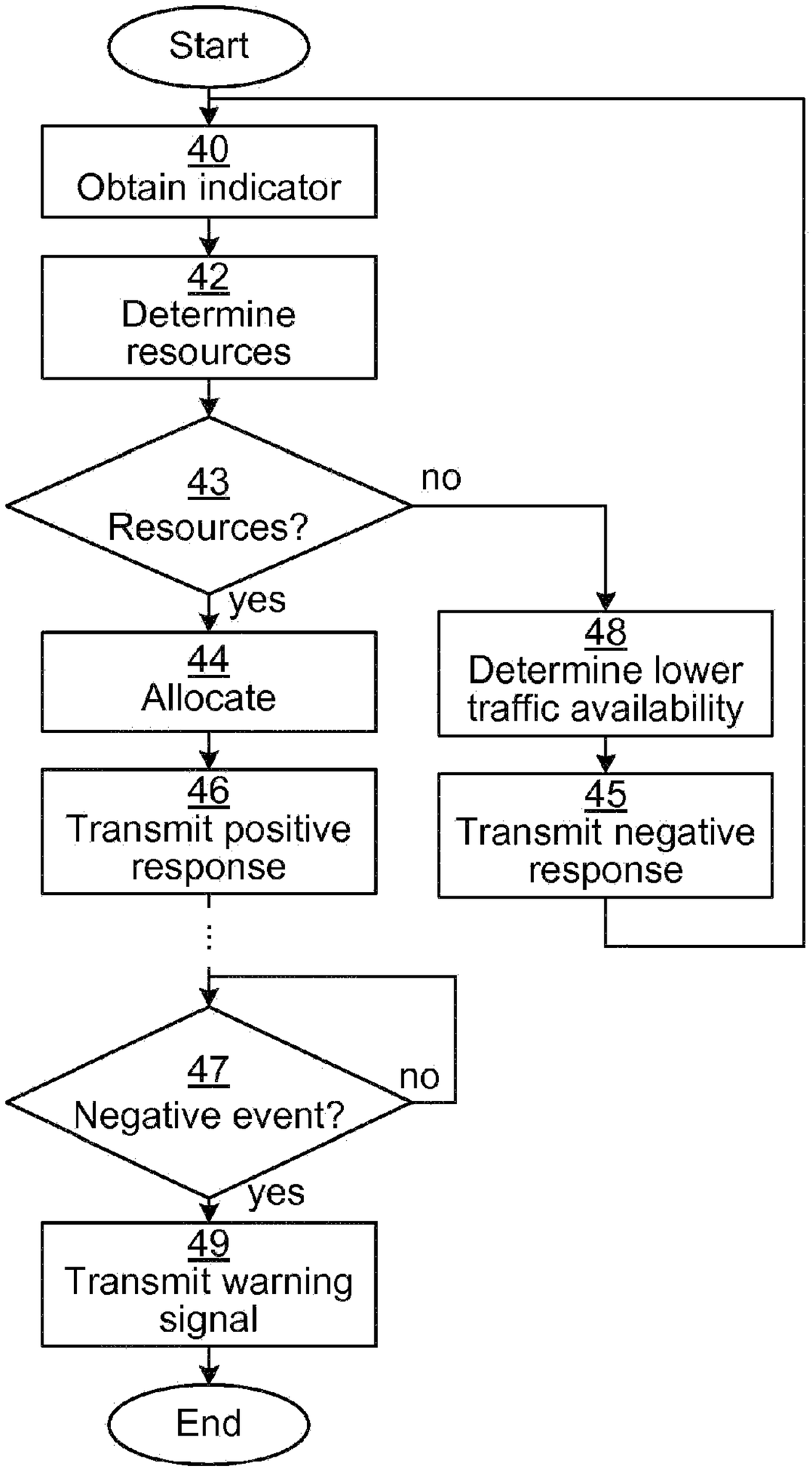
Start
40
Obtain indicator
42
Determine resources
43
Resources?
no
yes
44
Allocate
48
Determine lower traffic availability
46
Transmit positive response
45
Transmit negative response
47
Negative event?
no
yes
49
Transmit warning signal
End

Fig. 5B

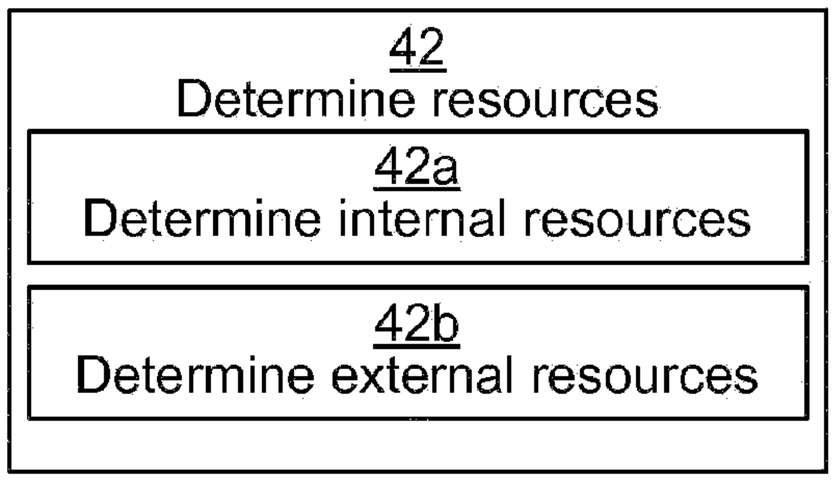

Fig. 5C

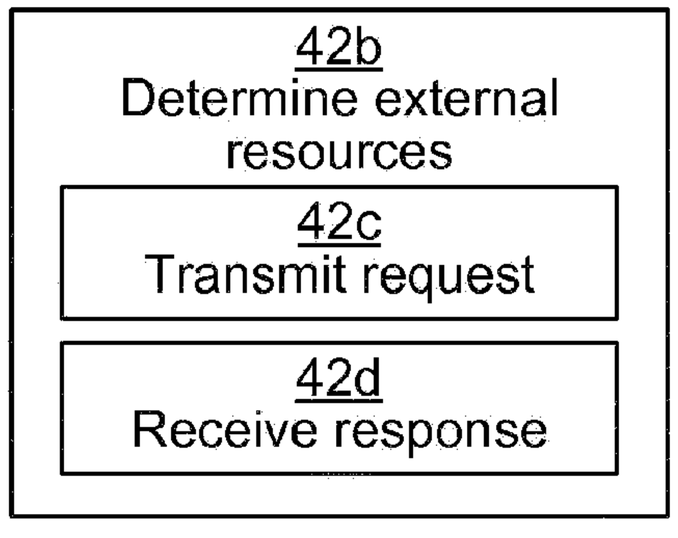

Fig. 5D

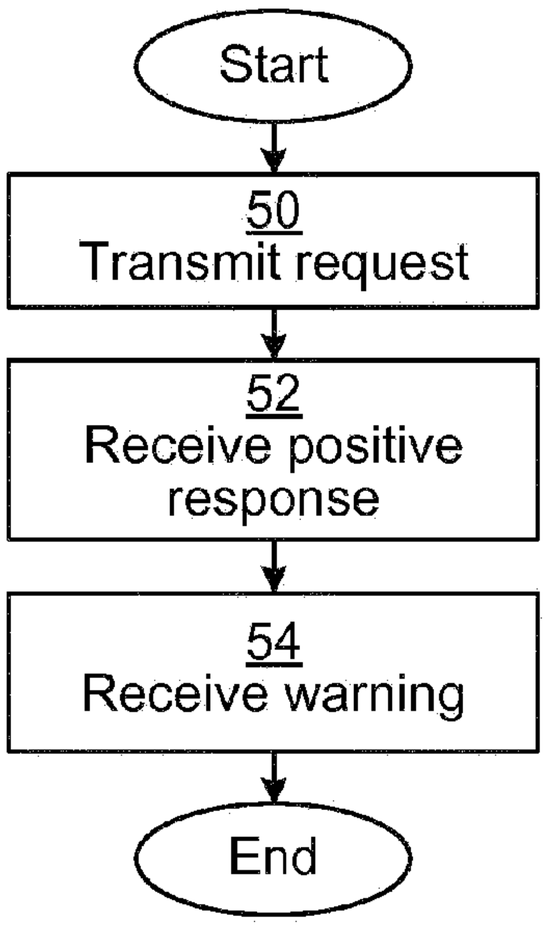

Fig. 6A

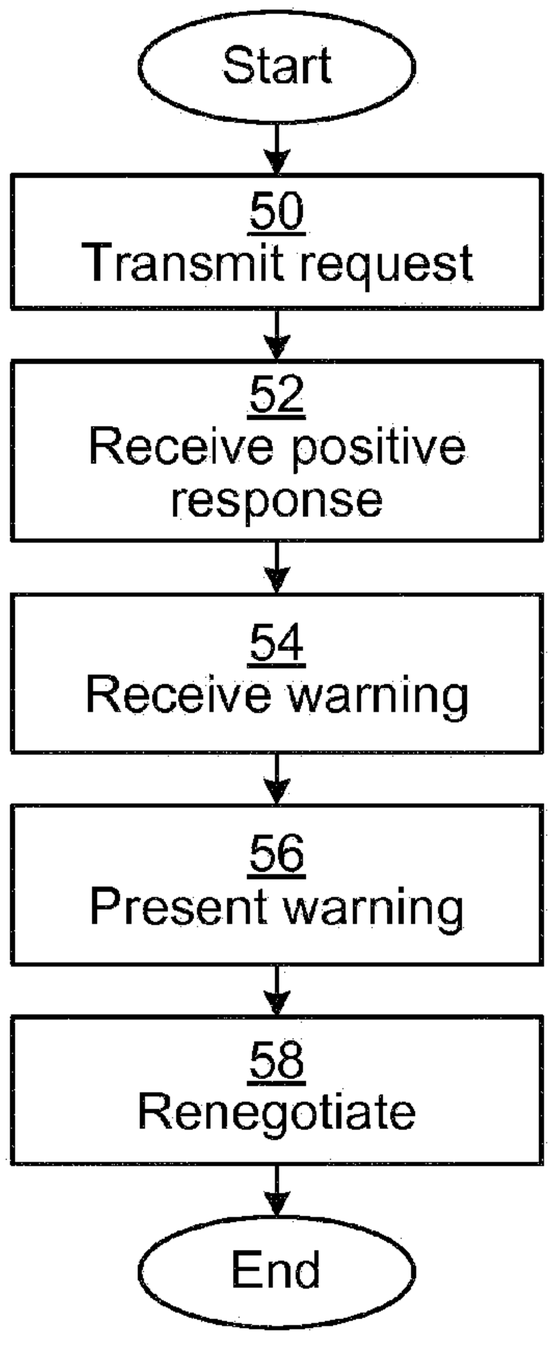

Fig. 6B

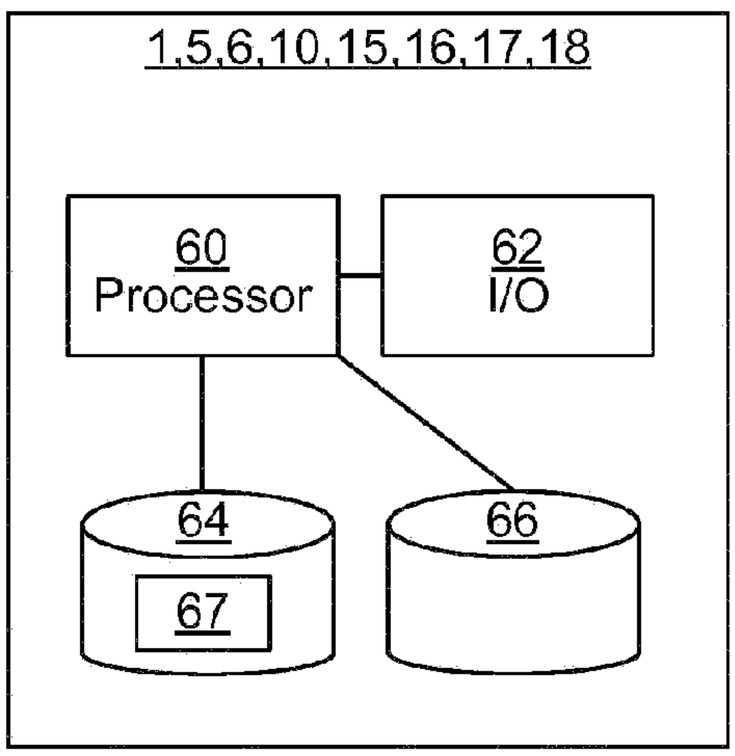

Fig. 7

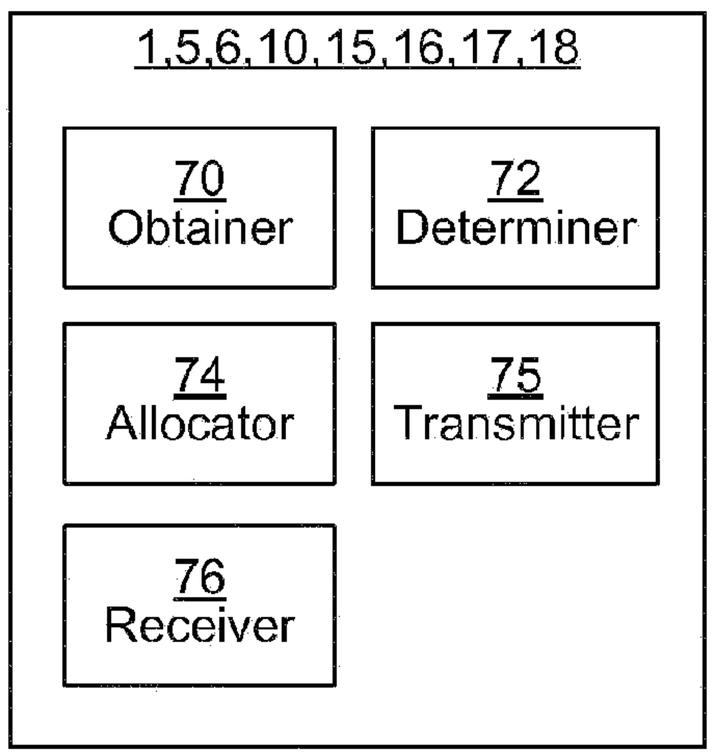

Fig. 8

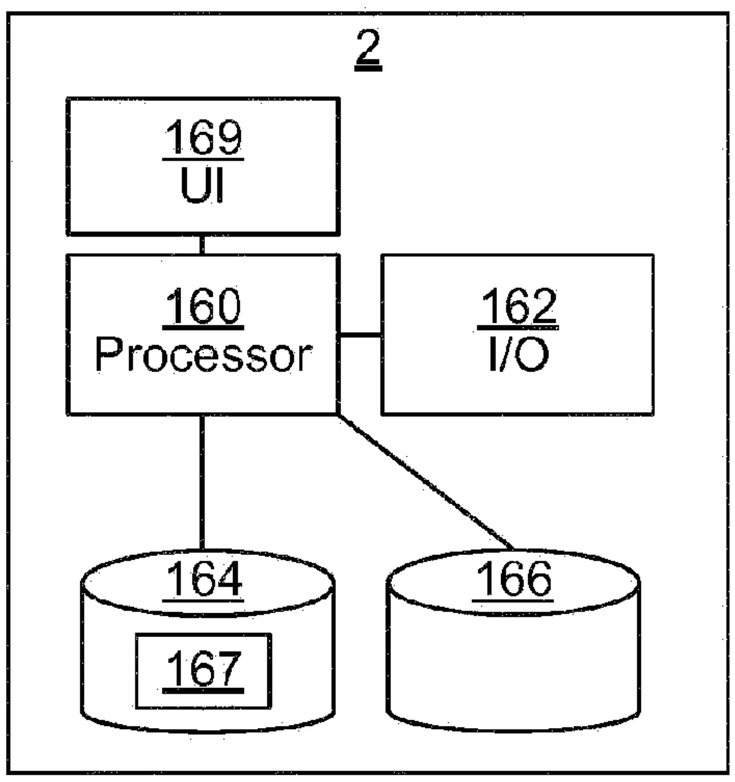

Fig. 9

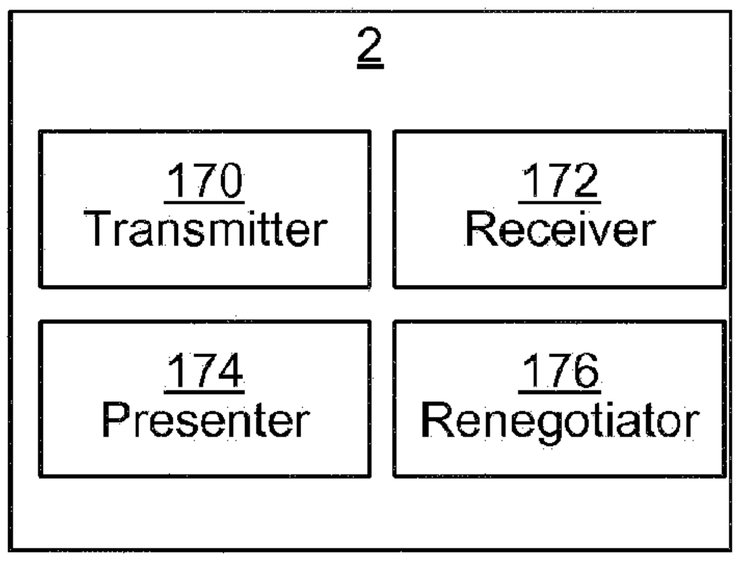

Fig. 10

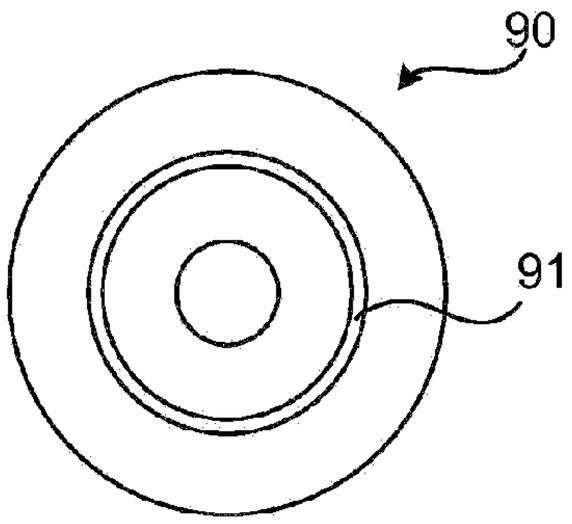

Fig. 11