

US009271183B1

(12) **United States Patent**
Oroskar et al.

(10) **Patent No.:** **US 9,271,183 B1**
(45) **Date of Patent:** **Feb. 23, 2016**

(54) **MANAGING ACCESS CHANNEL RESOURCES BY BUFFERING SERVICE REQUESTS**

FOREIGN PATENT DOCUMENTS

CN WO2011160287 * 12/2011 H04W 28/02

(71) Applicant: **Sprint Spectrum L.P.**, Overland Park, KS (US)

OTHER PUBLICATIONS

(72) Inventors: **Siddharth S. Oroskar**, Overland Park, KS (US); **Maulik K. Shah**, Overland Park, KS (US); **Jasinder P. Singh**, Olathe, KS (US)

WO2011160287A1 (with Machine Translation), Yijun Yu, Methods and Devices for Controlling Congestion, Jun. 22, 2010, p. 6+.*
LTE Quick Reference, Retry Test / Negative Test / Reject Test, http://www.sharetechnote.com/html/Handbook_LTE_RetryTest.html (retrieved Nov. 2013), 4 pages.

(73) Assignee: **Sprint Spectrum L.P.**, Overland Park, KS (US)

Rach, http://www.sharetechnote.com/html/RACH_LTE.html (retrieved Nov. 2013), 24 pages.

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 176 days.

Chris Johnson, Long Term Evolution (LTE) in Bullets 2nd Edition, 26.1 RRC Connection Establishment, available at <http://lte-bullets.com/LTE%20in%20Bullets%20-%20RRC%20Connection%20Establishment.pdf> (retrieved Nov. 2013), 5 pages.

(Continued)

(21) Appl. No.: **14/147,083**

Primary Examiner — Warner Wong

(22) Filed: **Jan. 3, 2014**

(51) **Int. Cl.**
H04W 28/02 (2009.01)
H04W 72/04 (2009.01)
H04W 72/12 (2009.01)

(57) **ABSTRACT**

(52) **U.S. Cl.**
CPC **H04W 28/0289** (2013.01); **H04W 72/0486** (2013.01); **H04W 72/1252** (2013.01)

A method for managing access channel traffic during congested network conditions is disclosed. A wireless connection device (WCD) sends a service request to a particular base station for a particular quantity of network resources. Upon the particular base station determining that it is too loaded with network traffic to accommodate the service request, the base station saves an indication of the WCD's connection request, and responds with a communication to instruct the WCD to idle and wait for a page from the base station. The WCD, upon receiving the communication, does not send a service request for a greater quantity of network resources than the particular quantity until the WCD receives a page. Once the base station determines it has sufficient network resources to allocate to the WCD, the base station sends a page to the WCD, which causes the WCD to send another service request.

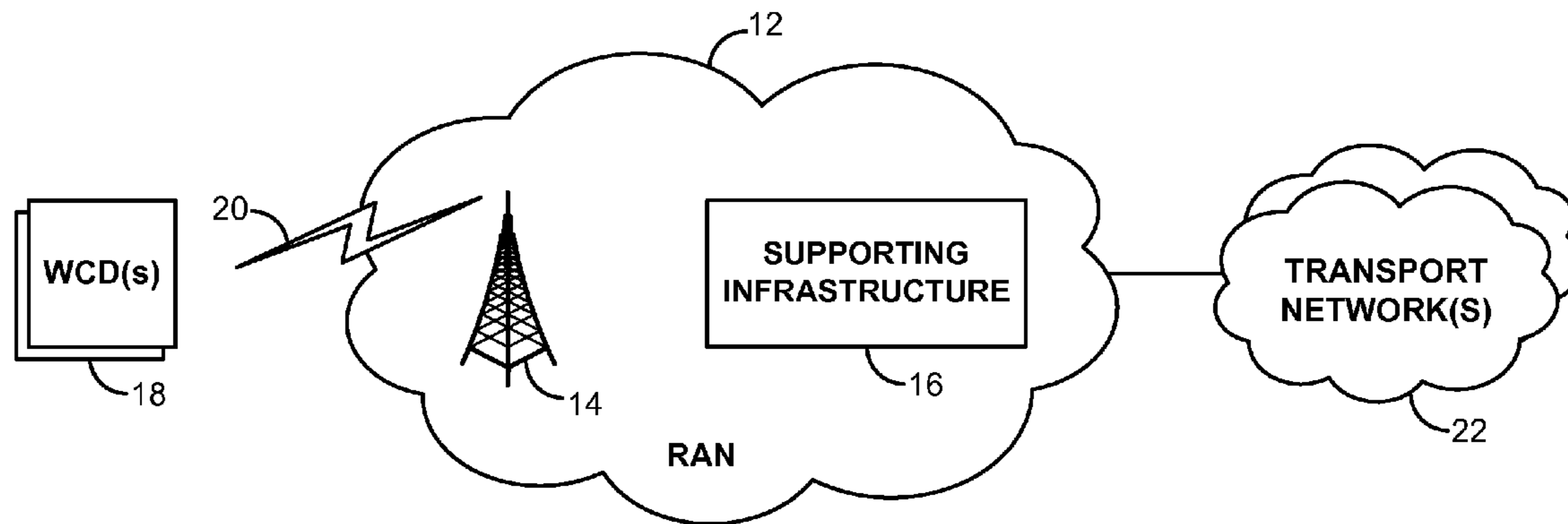
(58) **Field of Classification Search**
CPC . H04L 47/10; H04L 2012/5631; H04L 47/15; H04L 47/2441; H04L 47/70; H04L 2012/5629; H04L 49/255; H04W 28/02; H04W 28/0205; H04W 28/0247; H04W 28/0289; H04W 40/04; H04W 40/248
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

2011/0310857 A1* 12/2011 Vujcic 370/336

18 Claims, 6 Drawing Sheets



(56)

References Cited

OTHER PUBLICATIONS

Chris Johnson, Long Term Evolution (LTE) in Bullets, 23.1 RRC Connection Establishment, available at <http://lte-bullets.com/>

LTE%20in%20Bullets%20-%20RRC%20Establishment.pdf
(retrieved Nov. 2013), 4 pages.
Freescale Semiconductor, Inc. "Long Term Evolution Protocol Overview", Oct. 2008, 21 pages.

* cited by examiner

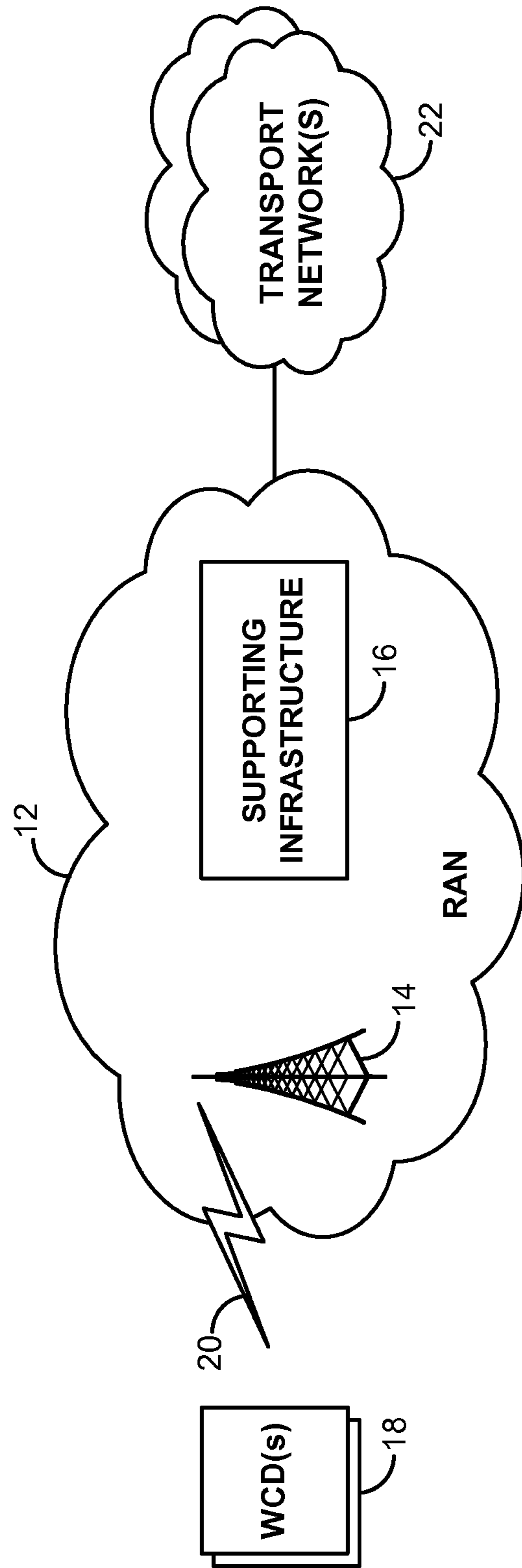


Fig. 1A

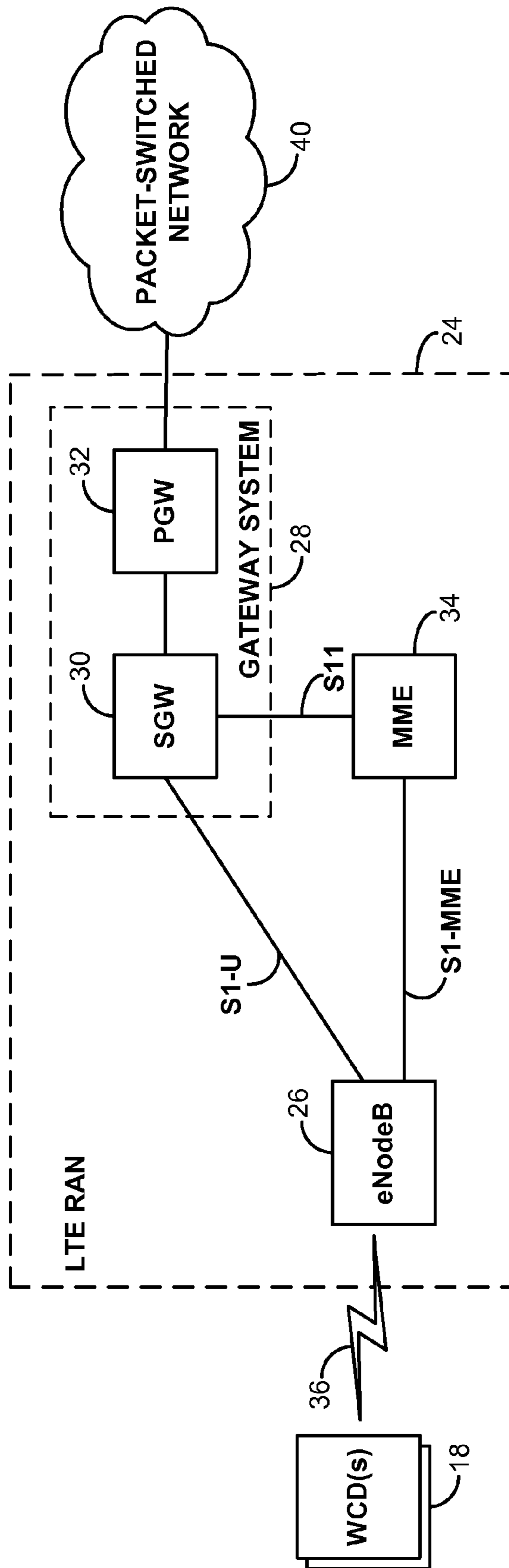


Fig. 1B

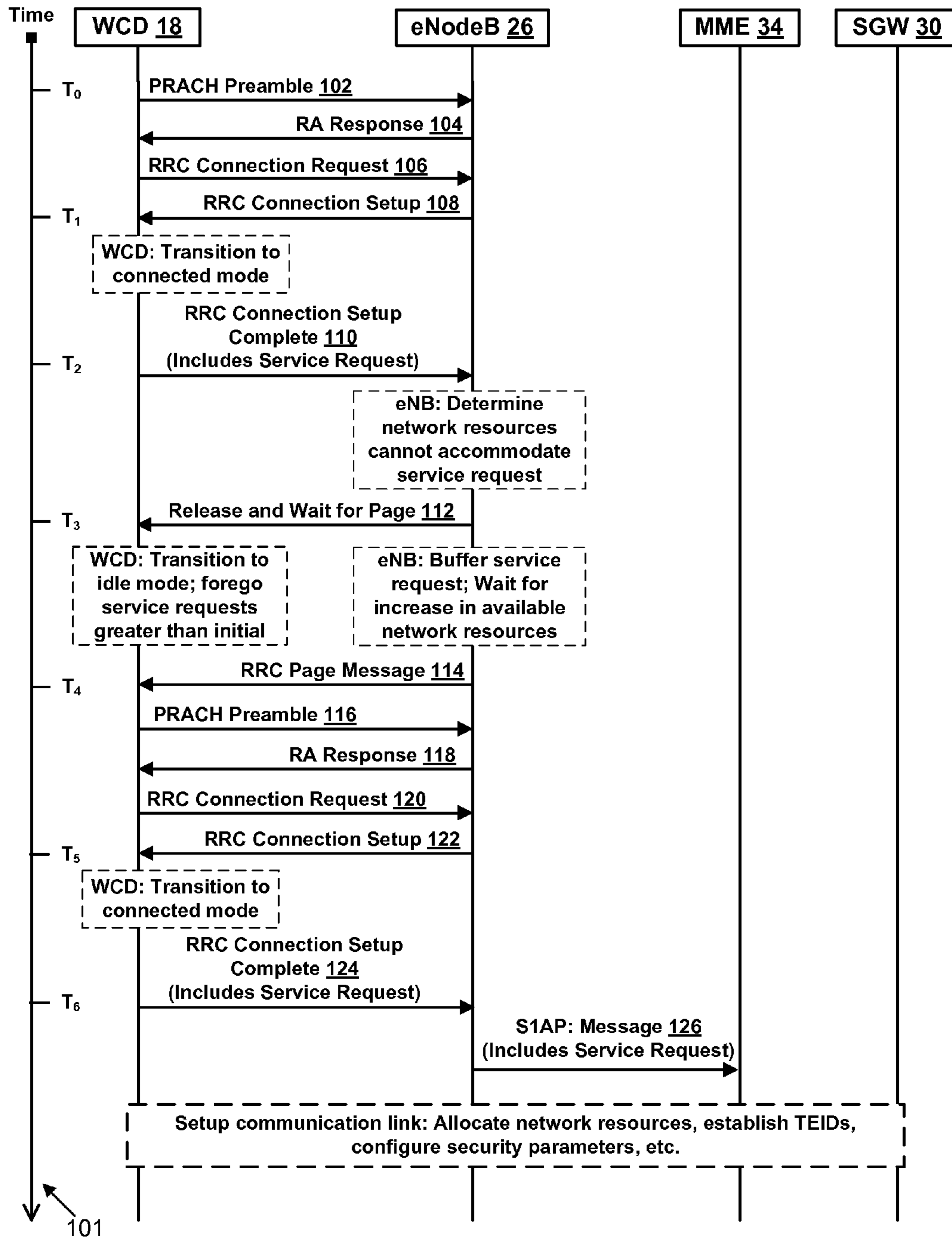


Fig. 2

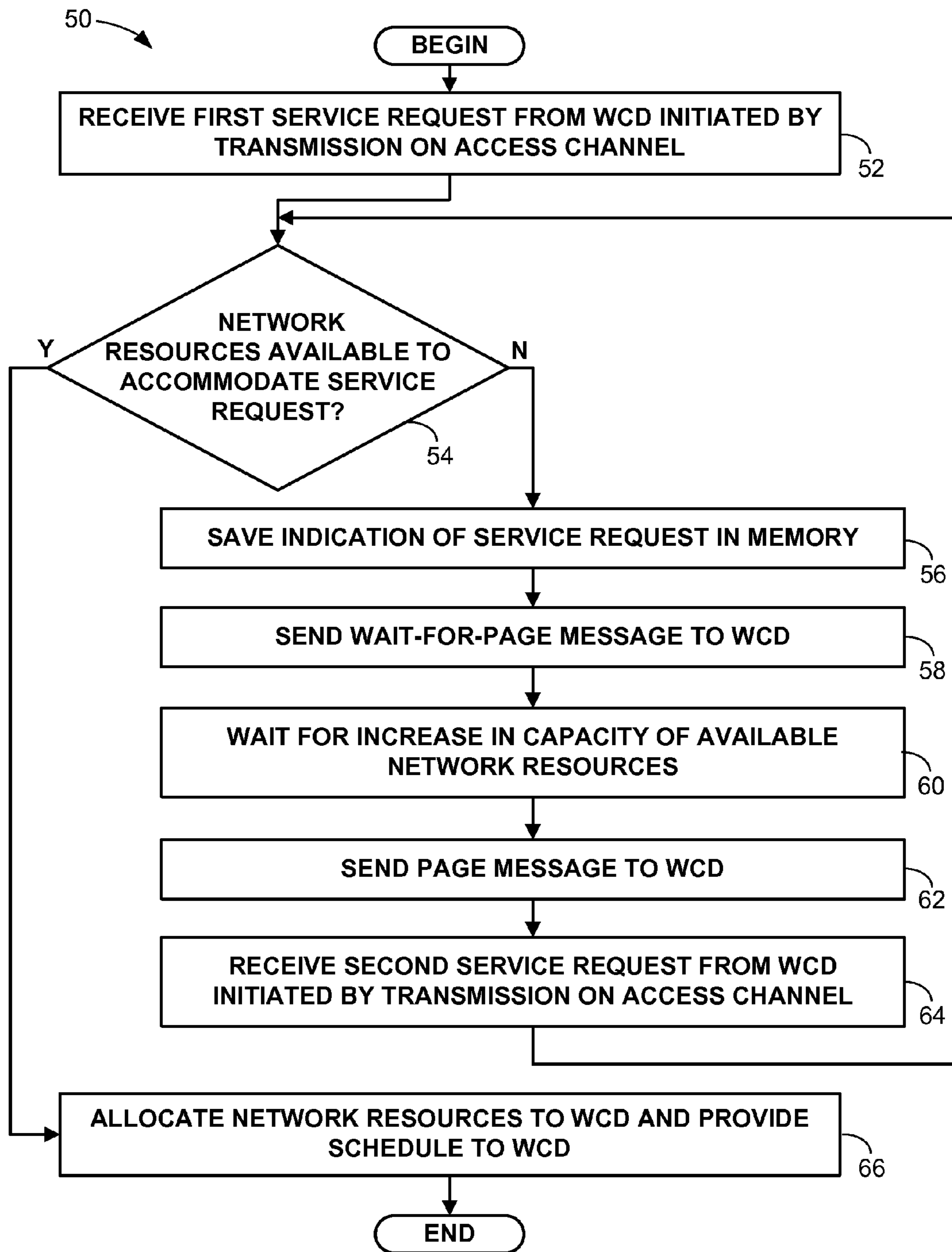


Fig. 3

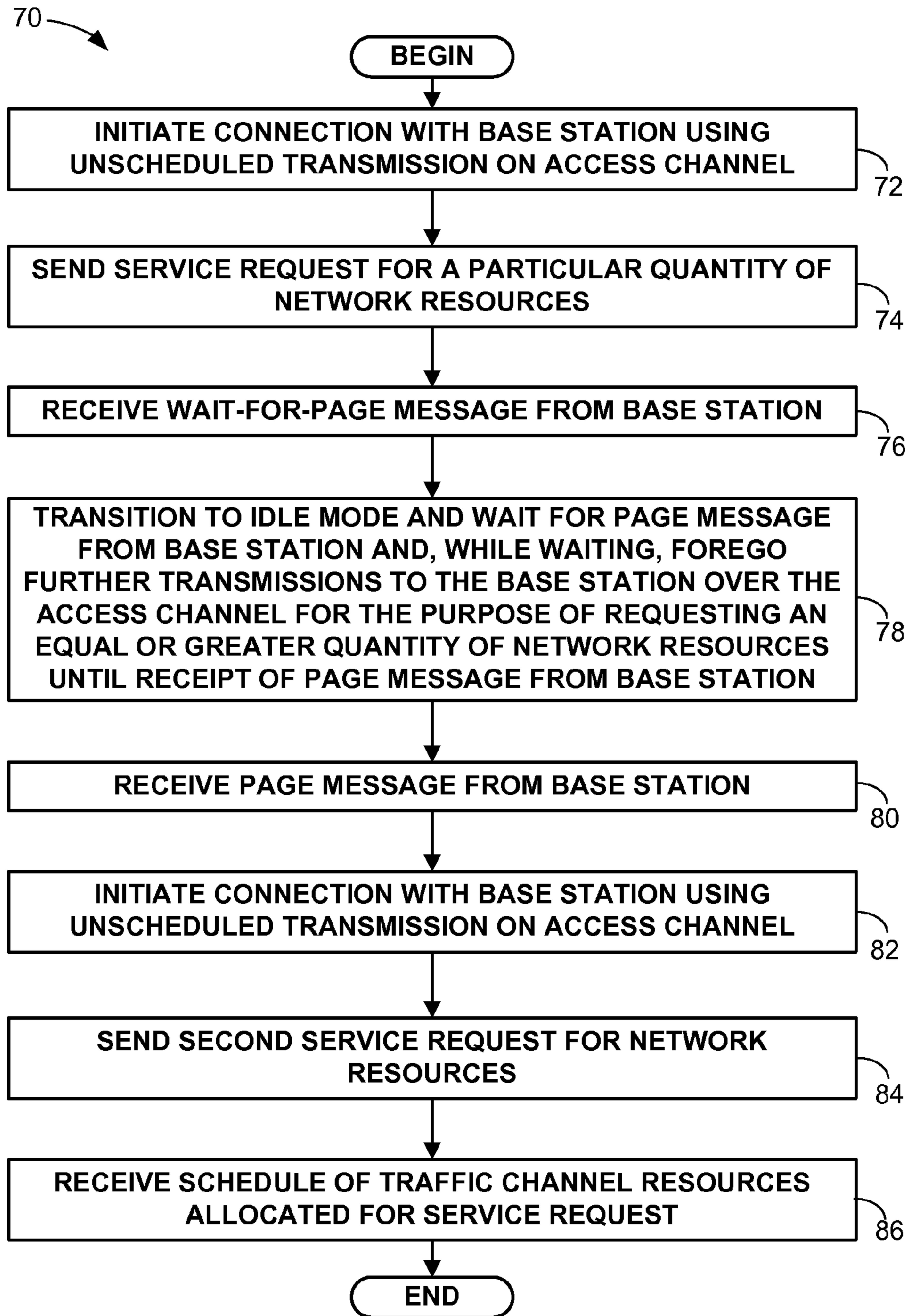


Fig. 4

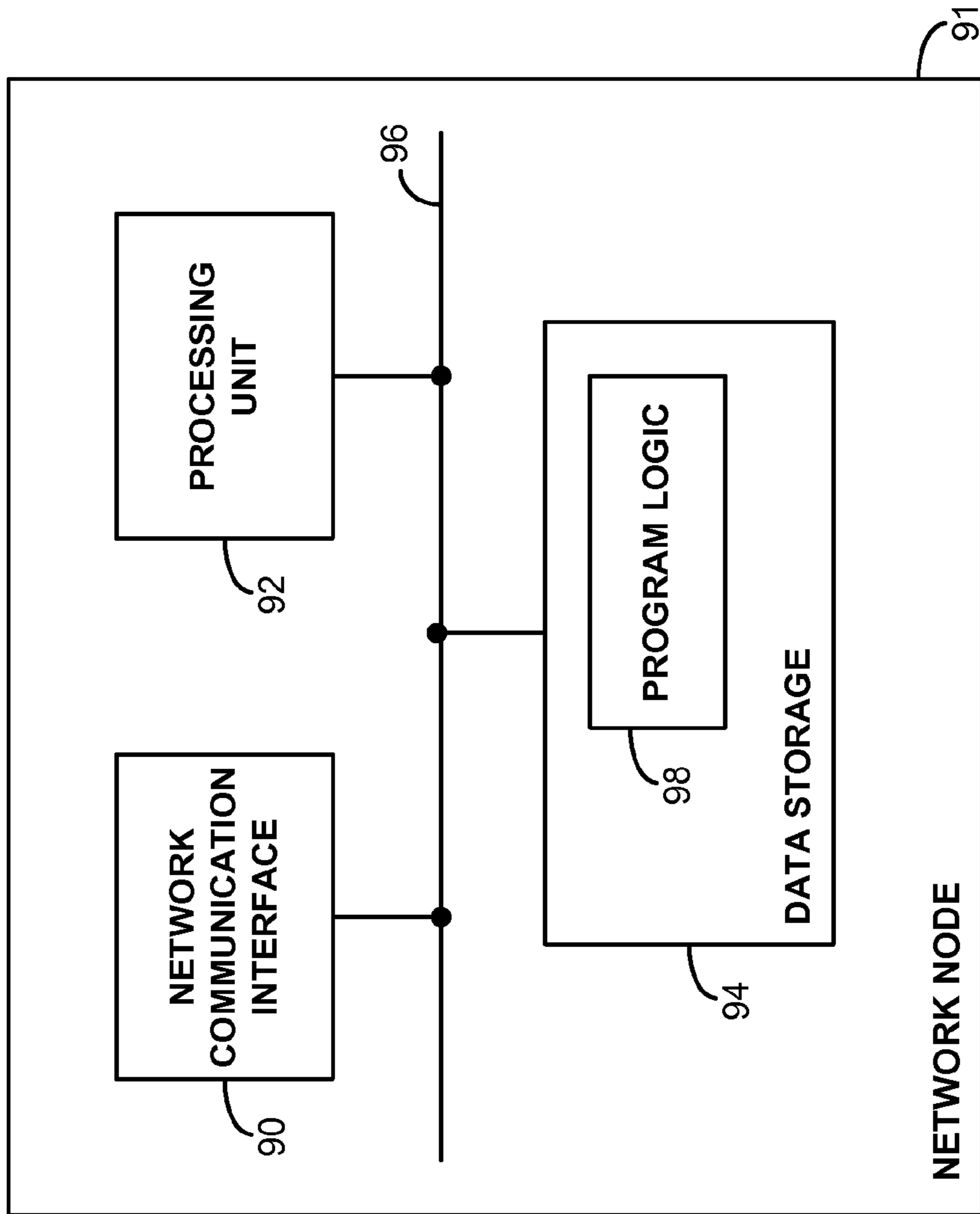


Fig. 5

**MANAGING ACCESS CHANNEL
RESOURCES BY BUFFERING SERVICE
REQUESTS**

BACKGROUND

Unless otherwise indicated herein, the materials described in this section are not prior art to the claims and are not admitted to be prior art by inclusion in this section.

A typical cellular wireless network system (wireless communication system) may include a number of base stations with antennas that radiate to define wireless coverage areas, such as cells and cell sectors. Within the wireless coverage areas, a subscriber (or user) accesses the communication services via a wireless communication device (WCD), which can communicate by exchanging radio frequency signals with the base stations. WCDs may include cell phones, tablet computers, tracking devices, embedded wireless modules, and other wirelessly equipped communication devices. In turn, each base station may be coupled with network infrastructure that provides connectivity with one or more communication networks, such as the public switched telephone network (PSTN) and/or a wide area network (WAN) for sending and receiving packet data (the internet, for instance). These (and possibly other) elements function collectively to form a Radio Access Network (RAN) of the wireless communication system. With this arrangement, a WCD within coverage of the RAN may communicate with various remote network entities.

In general, communications on the RAN are carried out in accordance with an air interface protocol that provides procedures for coordinating communications between the base stations and the WCDs. Examples of existing air interface protocols include, without limitation, Code Division Multiple Access (CDMA) (e.g., 1xRTT and 1xEV-DO), Long Term Evolution (LTE), Wireless Interoperability for Microwave Access (WiMAX), Global System for Mobile Communications (GSM), among other examples. Each protocol may define its own procedures for registration of WCDs, initiation of communications, handoff between coverage areas, and other functions related to air interface communication.

Protocols may also define procedures for managing communications from the base stations to the WCDs, which is referred to as a downlink, and for managing communications from the WCDs to the base stations, which is referred to as an uplink.

Depending on the specific underlying technologies and architecture of a given wireless communication system, the RAN elements may also take different forms. In a CDMA system configured to operate according to IS-2000 and IS-856 standards, for example, the antenna system is referred to as a base transceiver system (BTS), and is usually under the control of a base station controller (BSC). In a universal mobile telecommunications system (UMTS) configured to operate according to LTE standards, the base station is usually referred to as an eNodeB, and the entity that typically coordinates functionality between multiple eNodeBs is usually referred to as a mobility management entity (MME). In a CDMA system the WCD may be referred to as an access terminal (AT); in an LTE system the WCD may be referred to as user equipment (UE). Other architectures and operational configurations of a RAN are possible as well.

In accordance with the air interface protocol, each coverage area may operate on one or more carrier frequencies (or "carriers") and may define a number of air interface channels for conveying information between the base stations and the WCDs. These channels may be defined in various ways, such

as through frequency division multiplexing, time division multiplexing, and/or code-division multiplexing for instance.

By way of example, each coverage area may have a pilot channel, reference channel or other resource on which the base station may broadcast a pilot signal, reference signal, or other resource WCDs may detect as an indication of coverage and may measure to evaluate coverage strength. As another example, each coverage area may use an access channel, an uplink control channel, or other resource on which WCDs may transmit control messages such as connection requests and registration requests to the base station. And each coverage area may use a downlink control channel or other resource on which the base station may transmit control messages such as system information messages and page messages to WCDs. Each coverage area may then have one or more traffic channels or other resource for carrying communication traffic such as voice data, packet data, and/or other data between the base station and WCDs.

Furthermore, the available spectrum on the air interface can be divided into time-frequency segments to define the various channels, and also for allocation and scheduling purposes. In an LTE system, such time-frequency segments are commonly referred to as resource blocks and span 0.5 milliseconds in time and 180 kilohertz in bandwidth. For example, the base station can evaluate the demands for network resources amongst its served WCDs and then allocate its available resources to those WCDs to accommodate those demands to the extent sufficient resources are available. The control channels may be used to communicate between the WCDs and the base station to facilitate evaluation of the various network demands and also to notify the WCDs of their assigned network resources once resources are allocated.

While communications are generally scheduled by the base station, initiation of communication from a WCD to a base station generally involves at least one unscheduled transmission from the WCD to the base station. In accordance with the air interface protocol, the base stations may provide for a shared access channel on which unscheduled messages can be sent from WCDs to the base station to notify the base station of the WCD's presence in the base station's coverage area (e.g., for registration purposes). Once such an initial unscheduled transmission is received, the base station can then allocate initial uplink and/or downlink resources to communicate with the WCD as necessary to manage further communications. In particular, the WCD may use the initial uplink resources to send information regarding the quantity of network resources sought by the WCD. Such a communication from the WCD is referred to as a service request. The base station can then allocate sufficient network resources to accommodate the service request and send an indication of the allocation to the WCD.

However, in some cases a base station is too loaded with existing network traffic to allocate the resources sought by a WCD. In that case, after receiving an initial communication from the WCD over the access channel, the base station may respond by rejecting the WCD's attempt to connect. After receiving an indication of the base station's rejection, the WCD may then wait for some period before attempting another connection. In practice, after receiving an initial communication over the access channel, the base station may first allocate some initial control channel resources to receive further information from the WCD regarding the nature of the connection sought by the WCD, such as a service request. The base station can then, on the basis of that information, determine whether the base station's network resources are sufficient to accommodate the WCD and respond accordingly.

The access procedure described above, in which an unscheduled transmission is sent over the access channel, is used when a WCD does not already have any network resources, and therefore no means to send an uplink communication to the base station. In practice, this occurs primarily in two circumstances. First, the WCD may have data to transmit over the network, such as occurs when a call or internet session is originated on the WCD and the WCD begins buffering data to send out. Second, the network may receive data to communicate to the WCD, such as occurs when a remote entity initiates a call or other packet data communication addressed to the WCD. In the first case, the WCD may initiate the access procedure on its own in response to having data to send out. In the second case, the network first notifies the WCD that it should initiate the access procedure by sending the WCD a page message addressed to the WCD. Upon receiving the page message, the WCD initiates the access procedure to establish a connection with the network, at which point the data is delivered to the WCD.

As a result, WCDs without ongoing connections to the network continue to monitor particular downlink control channels that are used by the network to send out page messages. As used herein, WCDs with allocated network resources for ongoing communications are said to be operating in "connected mode." Connected WCDs are able to exchange data with remote entities over the network. On the other hand, WCDs without allocated network resources for ongoing communications are said to be operating in "idle mode." Idle WCDs monitor downlink control channels for pages and other system information, but do not generally transmit uplink communications back, which also results in reduced power consumption. To conserve network resources, WCDs may generally be configured to default to operating in idle mode, and transition to connected mode when data is ready to transmit or in response to receipt of a page message. Then, following a period of inactivity, the WCD can transition back to idle mode.

OVERVIEW

Disclosed herein is a process and corresponding system to manage communications over an access channel during loaded conditions. During a basic access procedure, if a base station is loaded with network traffic at the time it receives a service request, the base station may reject the service request and respond to the originating WCD to indicate the request was rejected. The communication indicating rejection may then cause the WCD to restart the access procedure and send another service request, perhaps after some delay. In the event the base station is still congested when the next service request arrives, the base station sends another rejection, and the WCD again initiates the access procedure and sends yet another service request. The cycle repeats until the base station has enough network resources to grant the service request. In the interim, the repeated unsuccessful attempts to employ the access procedure results in increased traffic on the access channel. Because the communications on the access channel are generally unscheduled, elevated traffic levels on the access channel exacerbates the possibility of interference between transmissions from separate WCDs.

The process disclosed herein avoids such repeated unsuccessful access channel communications during loaded conditions. According to the present disclosure, a base station that determines it is too loaded to accommodate a service request responds by instructing the originating WCD to idle and wait for a page message before submitting a service request for equal or greater network resources than sought by

the initial service request. Such a communication from the base station is referred to herein as a "wait-for-page" message. The WCD, upon receiving a wait-for-page communication, operates in idle mode and does not send an additional service request for an equal or greater quantity of network resources than sought initially. In effect then, the wait-for-page communication causes the WCD to forgo unscheduled transmissions on the access channel if the purpose of such transmissions is to submit a service request for an equal or greater quantity of network resources than sought by the initial service request.

After receiving the wait-for-page communication, and prior to receiving a page message, the WCD effectively acts as though it has been informed that the base station does not have enough network resources to grant service requests seeking an equal or greater quantity of network resources than the initial service request. Although the WCD may still be allowed to undergo the access procedure and to submit a service request for a quantity of network resources less than the quantity sought by the initial service request.

While the WCD idles and waits for a page message, the base station saves an indication of the service request and monitors its available network resources while waiting for an increase in available network resources. Upon determining that the available network resources have increased such that sufficient resources are available to accommodate the saved service request, the base station then sends a page message to the WCD to override the effects of the earlier wait-for-page communication. To facilitate such determinations, the base station may save, for each service request, an indication of a quantity of network resources sought by the service request. For instance, the base station may save an indication of a quality of service, a quantity of data, a desired latency, minimum bit rate, and/or other parameter(s) related to the quantity of network resources requested by a given service request. The base station can then use the saved data to make a subsequent determination that its available network resources are able to accommodate the service request saved in memory, and send a page message to the originating WCD.

After receiving the page message, the WCD initiates a connection with the base station (e.g., using an unscheduled transmission on the access channel) and submits another service request. The base station then grants the service request by allocating the requested network resources to the WCD. In addition, the base station coordinates with other network components to establish links to carry communications between the WCD and various remote networks in communication with the network. With the network links established, and resources allocated on the air interface, the WCD operates in connected mode to exchange data with remote entities over the network.

In some cases, a base station may maintain a buffer of service requests from different WCDs, and can page each WCD in the order in which the original service requests were received. For example, while loaded, the base station can save data indicative of each received service request. The saved data for each service request may include indications of the quantity of network resources requested, an identifier for the originating WCD, and the time of reception of the service request. The base station may then consider each saved service request, in the order of reception, and upon having sufficient resources to grant a next service request, send a page message to the corresponding WCD so as to cause that WCD to connect to the base station. As a result, the base station may page the WCDs in order, which helps to equalize the latency associated with establishing connections among different WCDs. By contrast, in a basic arrangement, previously

5

rejected WCDs might successfully establish connections based only on whichever WCD happens to submit a connection request immediately following an increase in available network resources.

Accordingly, in one respect, disclosed is a method of operating a wireless network system. The wireless network system can include a base station operating to serve wireless communication devices (WCDs) via one or more scheduled channels and an unscheduled access channel. The base station can be configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs. And a given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station can be configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel. The method can include receiving a first service request from a particular WCD. The first service request can specify a particular quantity of network resources. The method can include making a first determination that a capacity of available network resources on the scheduled channels is insufficient to allocate the particular quantity of network resources to the particular WCD responsive to receiving the first service request. The method can include, responsive to making the first determination, the base station: (i) sending a wait-for-page communication to the particular WCD, wherein the wait-for-page communication causes the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the particular quantity of network resources; (ii) storing data indicative of the received service request in a memory; (iii) waiting for the capacity of available network resources on the scheduled channels to increase; (iv) after waiting for the capacity of available network resources on the scheduled channels to increase, making a second determination that the capacity of available network resources on the scheduled channels is sufficient to allocate the particular quantity of network resources to the particular WCD; and (v) responsive to making the second determination, sending a page to the particular WCD.

In another respect, disclosed is a method of operating a wireless communication device (WCD) in a wireless network system. The wireless network system can include a base station operating to serve a plurality of WCDs including the WCD via one or more scheduled channels and an unscheduled access channel. The base station can be configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs. And a given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station can be configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel. The method can include sending a first service request from the WCD to the base station. The first service request can specify a particular quantity of network resources. The method can include receiving a wait-for-page communication from the base station responsive to sending the first service request. The method can include, the WCD, responsive to receiving the wait-for-page communication and prior to being paged, forgoing transmitting to the base station over the access channel for the purpose of initiating a service request specifying a

6

greater quantity of network resources than the particular quantity of network resources. The method can include receiving a page from the base station. The method can include sending a second service request to the base station responsive to receiving the page. The second service requests can be initiated by the WCD sending an unscheduled transmission over the access channel to the base station.

In another respect, disclosed is a wireless network system including a base station and a controller. The base station can have one or more antenna structures configured to wirelessly communicate with wireless communication devices (WCDs) served by the base station over one or more scheduled channels and over an unscheduled access channel. The base station can be configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs. And a given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station can be configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel. The controller can be operatively coupled to the antenna structures of the base station. The controller can be configured to receive, at the base station, a first service request from a particular WCD. The first service request can specify a particular quantity of network resources. The controller can be configured to, responsive to receiving the first connection request, make a first determination that a capacity of available network resources on the scheduled channels is insufficient to allocate the particular quantity of network resources to the particular WCD. The controller can be configured to, responsive to making the first determination: (i) send a wait-for-page communication to the particular WCD, wherein the wait-for-page communication causes the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the particular quantity of network resources; (ii) store data indicative of the received service request in a memory; (iii) wait for the capacity of available network resources on the scheduled channels to increase; (iv) after waiting for the capacity of available network resources on the scheduled channels to increase, make a second determination that the capacity of available network resources on the scheduled channels is sufficient to allocate the particular quantity of network resources to the particular WCD; and (v) responsive to making the second determination, send a page to the particular WCD.

Moreover, particular implementations of the present disclosure may include other examples for saving service requests, and/or other communications that involve unscheduled transmissions over an access channel, including wireless communications systems other than LTE systems.

These as well as other aspects, advantages, and alternatives will become apparent to those of ordinary skill in the art by reading the following detailed description, with reference where appropriate to the accompanying drawings. Further, it should be understood that the description provided in this overview section and elsewhere in this document is provided by way of example only.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A is a simplified block diagram of an example wireless communication system in which the present disclosure can be implemented.

FIG. 1B is a simplified block diagram of an example LTE system in which the present disclosure can be implemented.

FIG. 2 is a signal flow diagram illustrating signaling between network nodes that occurs when buffering a service request in an LTE system.

FIG. 3 is a flow chart depicting functions that can be carried out by a base station in a radio access network in accordance with an example method.

FIG. 4 is another flow chart depicting functions that can be carried out by a wireless communication device in accordance with an example method.

FIG. 5 is a simplified block diagram of a network node arranged to carry out various functions in accordance with the present disclosure.

DETAILED DESCRIPTION

Together with the drawings, the foregoing description provides an example wireless communication systems in which the present disclosure can be implemented. It should be understood, however, that this and other arrangements described herein are set forth only as examples. As such, those skilled in the art will appreciate that other arrangements and elements (e.g., machines, interfaces, functions, orders, and groupings of functions, etc.) can be used instead, and that some elements may be omitted altogether. Further, many of the elements described herein are functional entities that may be implemented as discrete or distributed components or in conjunction with other components, and in any suitable combination and location. Various functions described herein as being performed by one or more entities may be carried out by hardware, firmware, and/or software. For instance, various functions may be carried out by a processor executing instructions stored in memory.

I. Example Communication System Architecture

FIG. 1A depicts an example communication system that includes a radio access network (RAN) 12 having a representative base station 14 and supporting network infrastructure 16. The base station 14 includes antennas arranged to communicate with WCDs 18 in a coverage area over an air interface 20. RAN 12 then provides connectivity with one or more transport networks 22, such as the PSTN or the Internet for instance. With this arrangement, each WCD 18 that is in range of the base station 14 (i.e., in its coverage area) may register or attach with the RAN 12 and may engage in air interface communication with the base station 14 so as to communicate in turn with various remote entities on the transport network(s) 22 and/or with other WCDs served by the RAN 12.

In this arrangement, the air interface 20 may be configured according to a particular protocol, and the WCDs 18 and base station 14 may be programmed or otherwise configured to operate according to that protocol. According to the air interface protocol, the network resources provided by the air interface 20 may define time-frequency segments that divide the available spectrum for communications over the air interface 20. The base station 14 (alone or in coordination with supporting infrastructure 16 in the RAN 12) may allocate resources to the WCDs 18 in accordance with an allocation/scheduling routine based on various factors so as to accommodate the respective communication demands of the WCDs 18 and/or to achieve network performance criteria. The base station 14 can then provide each WCD with an indication of the schedule for their respective network resources, and the WCDs with allocated resources can then use their respective resources to exchange bearer data and control communications as necessary.

But a WCD without allocated resources for uplink communications initiates contact with the base station 14 by sending an unscheduled transmission over a shared access channel designated by the RAN 12. The unscheduled transmission may be used to initiate an access procedure, for example. Once the initial transmission is received, the base station 14 may assign initial resources for further communications between the WCD 18 and the base station 14. For example, the initial resources may be used for the WCD 18 to send the base station 14 a service request that specifies a particular quality of service (e.g., by specifying a minimum bit rate, latency, or another parameter) or otherwise designates a quantity of network resources sought by the WCD 18. Upon receiving the WCD's service request, the base station 14 may then determine whether to grant the service request, which determination may be based on whether the base station 14 has enough available network resources to accommodate the service request (e.g., to allocate the quantity of network resources sought by the service request).

According to the present disclosure, if the base station 14 determines it is too loaded by network traffic to grant the service request, the base station 14 can send a "wait-for-page" message to the WCD 18. The wait-for-page communication can cause the WCD 18 to transition to idle mode and to forgo further service requests for an equal or greater quantity of network resources until the WCD 18 receives a page message from the base station 14. While the WCD 18 waits, the base station 14 can save an indication of the service request, then wait for the base station's available network resources to increase. The saved indication of the service request may include, for example, data indicating the identity of the WCD 18 and the quantity of network resources sought.

After saving the service request, the base station 14 can monitor the increase in its available network resources and determine when its resources become sufficient to accommodate the saved service request. Upon determining the saved service request can be accommodated, the base station 14 can send a page message to the WCD 18. The page message overrides the effect of the earlier wait-for-page communication. Accordingly, the WCD 18 can then submit another service request to the base station 14, which involves another unscheduled transmission on the access channel. Because the second service request is sent in response to the base station 14 determining it has sufficient network resources available, the base station 14 is generally able to grant the second service request from the WCD 18, at which point the base station 14 can serve the WCD 18 in connected mode to facilitate further communications as necessary.

By contrast, as noted above, some protocols cause the base station 14 to send a rejection when the base station 14 is too congested to allocate resources. The rejection causes the WCD 18 to retry the random access procedure after some delay, at which point the base station 14 may still be too congested to allocate resources, which leads to repeated failed attempts to initiate connections using the access channel and exacerbates interference between the unscheduled transmissions. Thus, the presently disclosed enhancements, which allow for saving a service request until resources become available and then paging the WCD, obviates the unnecessary communications over the access channel while the base station 14 is loaded.

FIG. 1A generally represents any wireless communication system in which the present disclosure can be implemented, and, as noted above, variations on the arrangement shown are possible. To help illustrate features of the present disclosure, the remainder of this document will focus on an LTE system by way of example. Those of ordinary skill in the art will

readily appreciate, however, that the disclosed principles can be applied as well in other types of wireless communication systems in which connections are initiated using unscheduled transmissions from WCDs over an access channel, with variations where appropriate.

Example LTE Communication System

Accordingly, FIG. 1B is a simplified block diagram of a representative LTE system as an example of the system shown in FIG. 1A. As shown in FIG. 1B, the LTE system includes an LTE RAN 24 that primarily serves WCDs 18 with wireless packet data communication service (but may also provide voice call service, such as voice-over-IP service or circuit-switched fallback service). The LTE RAN 24 is shown including a representative LTE base station 26 known as an eNodeB, a gateway system 28 including a serving gateway (SGW) 30 and a packet data network (PDN) gateway (PGW) 32, and a mobility management entity (MME) 34.

In practice, the eNodeB 26 includes an antenna structure and associated equipment for engaging in LTE communication over an LTE air interface 36 with WCDs 18 to thereby provide the LTE RAN 24 with connectivity to the WCDs 18. The gateway system 28 provides the LTE RAN 24 with connectivity to packet-switched network(s) 40 such as the Internet and/or private networks. The various network elements in the LTE RAN 24 are linked together to facilitate communications between the served WCDs 18 and the remote packet-switched network 40. The eNodeB 26 communicates with the SGW 30 over an S1-U interface, the eNodeB 26 communicates with the MME 34 over an S1-MME interface, and the MME 34 communicates with the SGW 30 over an S11 interface. Although each of these interfaces is shown as a direct link in FIG. 1B, in practice the various elements of the LTE RAN 24 may sit as nodes on a wireless service provider's core packet network, and so these and other interfaces described herein may instead be logical connections over that packet network.

With this network arrangement, when a WCD 18 enters into coverage of eNodeB 26, the WCD 18 may scan for system information broadcast by the eNodeB 26. The WCD 18 can then use the system information to register or "attach" with the LTE RAN 24 by transmitting an attach request to the eNodeB 26, which the eNodeB 26 forwards to the MME 34. The MME 34 and other network elements may then carry out various functions to enable the LTE RAN 24 to serve the WCD 18. Among other functions, the MME 34 may add the WCD 18 to a database of registered WCDs managed by the MME 34, such as a home subscriber server (HSS)

The MME 34 may also communicate with the gateway system 28 and with the eNodeB 26 to setup and manage one or more bearer connections extending between the WCD 18 and the PGW 32 and thus between the WCD 18 and the packet-switched network 40. For instance, for each such bearer connection, the MME 34 may create and store in data storage a context record defining an evolved packet system (EPS) bearer identity for the WCD 18. The MME 34 may generate and transmit to the SGW 30 a create-session request identifying the serving eNodeB 26, which triggers setup of a tunnel between the SGW 30 and PGW 32 and assignment of an IP address for the WCD 18. In this process, the PGW 32 may also establish and store a context record for the WCD 18 and may assign an IP address to the WCD 18, and the PGW 32 may signal the assigned IP address in response to the SGW 30. The SGW 30 may then transmit to the MME 34 a create-session response specifying the assigned IP address.

Upon receipt of a create-session response from the SGW 30, the MME 34 may then further generate and transmit to the eNodeB 26 an attach-accept message identifying the SGW 30

and specifying the assigned IP address, which may trigger setup of a tunnel between the eNodeB 26 and the SGW 30 and assignment of a corresponding radio bearer identity defining a tunnel between the WCD 18 and the eNodeB 26. And the eNodeB 26 may forward to the WCD 18 an indication of the assigned IP address. Through this process, a bearer would thus be established between the WCD 18 and the PGW 32, including a series of tunnels extending (i) between the WCD 18 and the eNodeB 26, (ii) between the eNodeB 26 and the SGW 30, and (iii) between the SGW 30 and the PGW 32, and the WCD 18 would have an assigned IP address that it can use for communication on the packet-switched network 40. With each such bearer established, the eNodeB 26 may then serve the WCD 18 in connected mode in which the eNodeB 26 manages communications to and from the WCD 18 by allocating and scheduling resources for such communications.

Network Resource Allocation

As noted above, the LTE air interface 36 may span a particular segment of spectrum, with a given carrier frequency and bandwidth. The available spectrum may be subdivided into time-frequency segments referred to as resource blocks (RBs). For example, the RBs may each span 0.5 millisecond timeslot and have a bandwidth of 180 kHz. That is, during each 0.5 millisecond timeslot, the air interface may define a number of 180 kHz RBs. In practice, these RBs may be numbered sequentially across the bandwidth, and groups of the RBs may define RB groups that may be numbered as well. Overhead system information broadcast by the eNodeB 26 may inform the WCDs 18 in the coverage area which RB and RB groups are defined on the air interface. The WCDs can then use that information to understand RB assignments keyed to RB number and RB group number and otherwise to determine which RBs to use to send and receive data.

The eNodeB 26 (alone or in combination with MME 34 or other network infrastructure) may manage assignment of RBs for the WCD to use in receiving communications from the eNodeB 26 (i.e., downlink communications). For example, when a WCD 18 has data to transmit to the eNodeB 26 (i.e., uplink communications), the WCD 18 may transmit to the eNodeB 26 a service request for an assignment of uplink resources on the air interface 36. The service request may specify a quantity of data ready to be transmitted from the WCD 18, such as in a buffer status report. The service request may also specify a quality of service for the connection (e.g., a quality of service class identifier (QCI)), or may specify other aspects of the uplink assignment sought, such as target bit rate or latency, for instance. Additionally or alternatively, the service request may specify a particular bearer type for the air interface 36, such as in cases where the various RBs available for allocation are grouped to define different bearer types that may be used to carry different types of communications, for example, guaranteed bit-rate communications and non-guaranteed bit-rate communications and the like.

The eNodeB 26 may then perform a scheduling process in which the eNodeB 26 determines how many RBs to assign for uplink transmission to provide the requested type and/or quantity of data transmission requested by the WCD 18. In one approach, upon receiving a service request, the eNodeB 26 refers to a schedule of RBs in upcoming timeslot(s), and the eNodeB 26 assigns to the WCD 18 a number of RBs from those timeslot(s) as necessary and available. The eNodeB 26 may then transmit to the WCD 18 a control message (downlink control information) specifying the assigned RBs by RB number or RB group number, or otherwise indicating the schedule of network resources allocated to the WCD 18. In some cases, semi-persistent scheduling is employed such that a given allocation provides resources with a given quality

level on an ongoing basis until the allocation is changed. Semi-persistent scheduling can thereby reduce the number of control messages used to indicate assigned resources.

Upon receipt of an assignment of RBs, the WCD 18 may then transmit some or all of the data that prompted the initial service request as IP packets using the assigned RBs. Each packet bears the WCD's assigned IP address as source address and an appropriate IP address as destination address. Upon receipt of the packets, the eNodeB 26 may then forward the packets using an appropriate tunnel to the gateway system 28 for transmission by the PGW 32 to the packet-switched network 40, at which point the packets are routed according to their respective destination addresses.

Similar to the allocation of uplink resources, the eNodeB 26 (alone or in combination with associated network infrastructure) may manage assignment of RBs for the WCD 18 to receive communications from the eNodeB 26 (i.e., downlink communications). For example, when the LTE RAN 24 has data to communicate to the WCD 18 (e.g., due to a remote-initiated phone call, incoming email, etc. reaching the gateway system 28), the eNodeB 26 may assign one or more upcoming RBs on the downlink for use by the WCD 18. The eNodeB 26 can then inform the WCD 18 of the assigned RBs and transmit the data to the WCD 18 using the assigned RBs.

WCD Operation Modes

Once the WCD 18 is registered with the eNodeB 26, the WCD 18 can then operate in an idle mode in which the WCD 18 monitors a downlink control channel to receive overhead information and to check for any page messages pertinent to the WCD 18. While in idle mode, the WCD 18 is generally not engaged in ongoing communication using the LTE RAN 24. To conserve network resources while the WCD 18 is in idle mode, the WCD 18 may not sustain a continuous connection with the eNodeB 26, and may instead "wake up" as needed to activate its radio and monitor for relevant communications over the downlink control channel. Accordingly, the downlink control channel may be configured to facilitate discontinuous transmission from the eNodeB 26. Generally, while in the idle mode, the WCD 18 does not have allocated uplink resources with which to send communications back to the eNodeB 26.

To engage in communication using the RAN, the WCD 18 first transitions to a connected mode, active mode, or other mode in which the eNodeB 26 allocates network resources for the WCD 18 to use. When the LTE RAN 24 has a communication (such as a voice call or other traffic) to provide to a WCD 18 that is registered with the network but is operating in the idle mode, the network may page the WCD 18 in an effort to facilitate assigning traffic channel resources to use to deliver the data to the WCD 18. The WCD 18 receives the page message, and initiates connection with the eNodeB 26. The eNodeB 26 and/or other network entities may then establish links between the eNodeB 26 and other network entities to carry communications to and from the WCD 18 over the LTE RAN 24. The eNodeB 26 also assigns resources on the air interface 36, and thus transitions the WCD 18 to a connected mode in which the WCD 18 can engage in communication with remote entities on the packet-switched network 40.

Likewise, when the WCD 18 is idle and seeks to initiate (or originate) a communication (such as to place a voice call or engage in other bearer communication), the WCD 18 may initiate connection with the eNodeB 26 and send a request for a desired quantity of network resources to the eNodeB 26. Links for carrying bearer data to and from the WCD 18 can then be established by the network, and the eNodeB 26 can assign network resources to the WCD 18 for use to carry

communications over the air interface 36, similarly transitioning the WCD 18 to a connected mode in which the WCD 18 can engage in ongoing communication over the LTE RAN 24.

The eNodeB 26 (and/or other components in the LTE RAN 24) can also cause the WCD 18 to transition from connected mode back to idle mode. For example, following a period of inactivity by the WCD 18, the network can send the WCD 18 a release message, which causes the network to tear down the network links established for the WCD 18, and to cease allocating resources to the WCD 18. After returning to idle mode, the WCD 18 continues to monitor downlink control channels for page messages and system information, but does not have resources allocated for sending messages back to the network.

To track the various WCDs in connected/idle modes associated with a given eNodeB, the MME 34 may assign a state to each WCD within its tracking network to reflect the current status of the WCDs as operating in connected mode or operating in idle mode. Based on the states associated with each WCD, the LTE RAN 24 then allocates resources to particular WCDs on an ongoing basis. At the same time, the individual WCDs may also have a state parameter that reflects the current state of the WCDs and determines the behavior of the WCDs. When transitioning between modes, the values of the state parameters maintained by the individual WCDs and the MME 34 can be adjusted so as to agree with one another. Although, in general, one state parameter may shift before the other depending on the manner in which the transition is initiated.

Random Access Procedure

While communications to and from the WCD 18 are generally scheduled by the eNodeB 26 as noted above, initiation of communication from the WCD 18 to the eNodeB 26 generally involves at least one unscheduled transmission from the WCD 18 to the base station 26. In particular, while the WCD 18 is in idle mode, the WCD 18 generally does not have resources allocated for sending an uplink communication. In the LTE RAN 24, the idle WCD 18 may initiate a connection with the eNodeB 26 using a contention-based random access procedure. The procedure involves contention between competing WCDs, because unscheduled transmissions from different WCDs may interfere with one another, given that the eNodeB 26 does not coordinate/schedule those transmissions. Two interfering WCDs are then said to be in contention for the attention of the eNodeB 26, and the process to determine which of the two WCDs (if any) is successful is referred to as resolving the contention.

The random access procedure begins with the eNodeB 26 broadcasting system information that specifies its physical random access channel (PRACH) for carrying unscheduled transmissions. For example, the system information may designate time-frequency windows (e.g., RBs) designated for PRACH transmissions as well as other network information (e.g., network capabilities, synchronization information, etc.) The unscheduled transmissions can each include a preamble code to differentiate transmissions received in the same time-frequency window. WCDs in the base station's coverage area detect the system information broadcast, randomly select a preamble code, and send a message using the selected preamble code to the base station over PRACH, which message is referred to as a preamble message. The preamble codes are configured such that the eNodeB 26 is able to use code differentiation to resolve preamble messages with different preamble codes. But, because the preamble codes are independently randomly selected by the WCDs 18, it is possible for multiple WCDs to select the same preamble code, in which

case the messages interfere with one another and result in the previously noted contention. If one of the interfering preamble messages is stronger than the others, the eNodeB 26 may be able to successfully decode the message with the strongest signal despite the increased noise from the others. But if none of the interfering preamble messages are significantly stronger than the others, the eNodeB 26 may fail to decode any of them.

The eNodeB 26 receives the preamble message over PRACH and responds with a random access response (RAR) during a response window that is based on the time-frequency window used to send the initial preamble message. As a result, all WCDs that sent preamble messages during a given time-frequency window monitor the same response window. A given RAR may therefore include data addressed to each of the preamble codes received in the preceding PRACH time-frequency window. For example, the RAR may include indicators for each of the preamble codes the eNodeB 26 received, and the WCD(s) that used those codes for their preamble messages then use the indicators to identify the data addressed to them. If a given WCD does not identify an indicator corresponding to the preamble code used by that WCD (such as may occur if the preamble message could not be decoded by the eNodeB due to interference or otherwise), the WCD restarts the random access procedure by selecting a new preamble code and transmitting another preamble message over PRACH.

For a WCD that does identify an indicator corresponding to their selected preamble code in the RAR, the data addressed to the WCD may include an assignment of initial control channel uplink and downlink resources to allow the WCD to exchange further information with the eNodeB 26. In addition, the RAR can assign a temporary cell radio network temporary identifier (temporary C-RNTI) to identify the WCD 18 in subsequent communications, and can also indicate a timing offset for the WCD 18 to use to account for propagation delays to the eNodeB 26 or to otherwise synchronize uplink communications.

The WCD 18 uses the uplink resources assigned in the RAR and sends the eNodeB 26 a connection request. Among other information, the connection request may include a unique identifier for the WCD 18 (e.g., a temporary mobile subscriber identity (TMSI), a previously assigned C-RNTI, a random number, or another unique identifier). The WCD 18 then monitors the downlink resource granted in the RAR for a response from the eNodeB 26. The eNodeB 26 responds with a contention resolution message that includes the WCD's unique identifier. Receiving a matched unique identifier in the contention resolution message provides confirmation that the WCD's connection request was received by the eNodeB 26. A failure to match may occur, for example, if multiple WCDs transmitted the initial preamble message in the same time-frequency window using the same preamble code. If the signal from one of the WCDs is stronger than the other(s), the base station may be able to decode the preamble message (and subsequent connection request) from that one WCD, but not the other(s). The WCD that sent the weaker signal may still receive the RAR, but the unique identifier included in the connection request message (and subsequent response from the base station) allows a given WCD to determine whether the eNodeB 26 is receiving and decoding signals from that WCD or from another WCD.

The contention resolution message thus resolves the contention due to the possibility of multiple WCDs using the same preamble code. If a WCD receives a contention resolution message that does not match its unique identifier, the WCD discards its temporary C-RNTI and restarts the random

access procedure by selecting a new preamble code and transmitting another preamble message over PRACH. The WCD that passed contention resolution (i.e., that received back its matching unique identifier) promotes its temporary C-RNTI to a full C-RNTI for use in subsequent communications with the eNodeB 26, and continues exchanging messages with the base station as necessary to setup a connection.

In some arrangements, when the eNodeB 26 does not have enough available network resources to accommodate additional traffic, the eNodeB 26 may respond to a PRACH preamble message, or to a subsequent service request, by denying the attempt to connect and then sending an appropriate message to the originating WCD. To effect a denial, the eNodeB 26 may send a rejection message, a release message, and/or a redirection message. A rejection message may be sent in response to a connection request and is used to notify the WCD 18 that the eNodeB 26 is too congested by network traffic to allocate resources for the WCD 18. Instead, the rejection message instructs the WCD 18 to wait for some interval, in idle mode, until attempting another connection. A release message may be sent following a service request and, like the rejection message, causes the WCD 18 to transition to idle mode and wait for some interval before attempting another connection. A redirection message identifies another nearby eNodeB or other resource so as to cause the WCD 18 to scan for the other resource and initiate a connection, which thereby redirects the WCD 18 to the other identified resource.

Following a rejection message or release message, the idle WCD 18 generally initiates the random access procedure yet again to make a connection with the eNodeB 26. But the eNodeB 26 may still be congested when the next preamble message is received, at which point the eNodeB 26 may deny the WCD 18 yet again. The process of unsuccessfully undertaking the random access procedure may therefore be repeated until the eNodeB 26 has adequate network resources to allocate resources to the WCD 18, at which point a connection may be successfully established. In the interim, the communications between the WCD 18 and the eNodeB 26 consume resources on PRACH and various control channels. Among other effects, the elevated communications on PRACH leads to increased contention in the random access procedure. Embodiments of the present disclosure provide an enhanced approach to managing PRACH resources under loaded conditions. In practice, upon receiving a service request that cannot be accommodated due to network loading, rather than send a release message, the eNodeB 26 sends a message to instruct the WCD 18 to idle and wait for a page message before re-attempting the random access procedure for the purpose of submitting a service request for the same or greater quantity of network resources. Such a message is referred to herein as a wait-for-page communication. The eNodeB 26 saves the service request and waits for its capacity of available network resources to increase. Once sufficient network resources are available to accommodate the saved service request, the eNodeB 26 pages the WCD 18, which causes the WCD 18 to initiate the random access procedure, transition to connected mode, and submit another service request to the eNodeB 26.

II. Example Signal Flow

FIG. 2 shows an example signaling process between the WCD 18, the eNodeB 26, and the MME 34. In particular, FIG. 2 illustrates signaling that occurs in an LTE system when the wait-for-page communication is sent in response to a service request. A timeline 101 indicates the timing of certain signals in the signal flow diagram in FIG. 2. For example purposes, the various communications represented in FIG. 2 are designated in accordance with messages used in the radio

resource control (RRC) protocol, such as connection requests, service requests, acknowledgements, and so on. Moreover, it is noted that alternative implementations may include more or less specific communications between various entities, such as overhead communications, acknowledgements, status inquiries and reports, and so forth. The signaling process illustrated by FIG. 2 may therefore be supplemented and/or modified to comply with specific signaling protocols, standards, etc., as will be appreciated.

The signaling process begins with the WCD 18 initiating the random access procedure. In practice, the WCD 18 detects system information broadcast from the eNodeB 26, randomly selects a preamble code, and transmits a preamble message 102 to the eNodeB 26 over PRACH at time T_0 . The eNodeB 26 receives the preamble message 102 and responds with a random access response message (RAR) 104. The RAR 104 can assign initial uplink and downlink control channel resources to allow the WCD 18 to exchange further information with the eNodeB 26.

The WCD 18 uses the uplink resources assigned by the RAR 104 and sends the eNodeB 26 an RRC Connection Request 106. The WCD 18 then monitors the downlink resource indicated in the RAR 104 for an RRC Connection Setup 108 from the eNodeB 26. The WCD 18 receives the RRC Connection Setup 108 from the eNodeB 26, at time T_1 , which resolves the contention in the random access procedure (e.g., by matching a unique identifier included in the RRC Connection Request 106). The RRC Connection Setup 108 can also include a scheduling command for the WCD 18 to communicate additional setup information. In response to receiving the RRC Connection Setup 108, the WCD 18 transitions to connected mode, and sends an RRC Connection Setup Complete 110 to the eNodeB 26, which includes a service request (e.g., by including a buffer status report or the like). The eNodeB 26 receives the RRC Connection Setup Complete 110 at time T_2 , and determines whether to grant the service request. The service request may specify a quantity of data ready to be transmitted from the WCD 18, and may also specify a particular quality of service and/or bearer type for carrying that data. If the eNodeB 26 determines that sufficient network resources are available to accommodate the service request, the eNodeB 26 can communicate with the MME 34 and/or other network entities to establish tunnels for bearer data and allocate resources for communications on the air interface, as described above.

The eNodeB 26 can determine whether it has sufficient network resources available to accommodate an incoming service request by, for example, determining a quantity of resource blocks (RBs) necessary to accommodate the service request, and comparing the determined quantity with a quantity of available RBs in upcoming timeslots. The determination may additionally or alternatively be made based on, for example, comparing the quantity of available RBs with a threshold value. As used herein, available RBs refer to RBs in the eNodeB's designated spectrum that are not already allocated for use by another WCD. Thus, a given eNodeB's capacity of available network resources (e.g., quantity of available RBs) is generally a function of both the total bandwidth of that eNodeB's air interface and the network loading conditions on the eNodeB. In addition, the quantity of network resources available at any given time and/or requested by a given service request may be specified by a quality of service (QoS) parameter (e.g., a measure of bit rate, latency, etc.), a quality of service class identifier (QCI), a number of RBs (a bulk number or a number per unit time), a bearer type, or another measure. As a result, the determination of whether to grant a given service request may be made on the basis of

whether a particular number of RBs are available per unit time on an ongoing basis in addition to, or as an alternative to, whether a particular bulk number of RBs are available during one or more upcoming timeslots.

If the eNodeB 26 determines it does not have sufficient available resources to accommodate the service request (as in FIG. 2), the eNodeB 26 may send the WCD 18 a "wait-for-page" message 112 at time T_3 . The wait-for-page communication 112 is interpreted by the WCD 18 as a directive that the WCD 18 should, in addition to transitioning to idle mode, forgo submitting certain service requests until the WCD 18 receives a page message from the eNodeB 26. In particular, the wait-for-page communication 112 directs the WCD 18 to idle and to not send service requests for a quantity of network resources equal or greater than the quantity sought by the initial service request (in the RRC Connection Setup Complete 110). As such, the WCD 18 is directed to not re-attempt the random access procedure (and thus to refrain from transmitting over PRACH) for the purpose of submitting such a service request until the WCD 18 is paged by the eNodeB 26. Thus, after receiving the wait-for-page communication 112, the WCD 18 can be configured to operate in idle mode and, while idling, to consider any proposed service requests before initiating the random access procedure for the purpose of submitting those service requests. In practice, the WCD 18 can check whether a proposed service request specifies an equal or greater quantity of network resources than the quantity sought initially, and to only proceed with the random access procedure if the proposed service request is for a lesser quantity. While idling, the WCD 18 also monitors a predetermined downlink control channel (e.g., the PDCCH) for page messages from the eNodeB 26 and other system information.

To facilitate clarity and expediency in the remaining description herein, reference is made to WCDs that forgo service requests for a "greater quantity" of network resources than an initial service request. But this description encompasses forgoing service requests for an equal or greater quantity of network resources. For instance, in network resources may be allocated in distinct quantities or RB groups associated with different bearer types, or traffic types, and each the various types may be sorted according to a hierarchy based on the quantity of network resources consumed by each. As such, WCDs may then forgo service requests that seek a bearer type of an equal or greater hierarchical level than the initial service request.

Moreover, in some examples, rather than directing the WCD to forgo PRACH transmissions only to the extent they are used to initiate a service request for a greater quantity of RBs than originally sought, the wait-for-page communication 112 may simply direct the WCD 18 to forgo all PRACH transmissions. Alternatively, the wait-for-page communication 112 may itself include an indication of a maximum quantity of network resources that may be requested by a service request. In such an example, the wait-for-page communication 112 may be interpreted as a directive that prohibits PRACH transmissions to the extent they are used to initiate service requests for network resources in excess of the maximum indicated.

The wait-for-page communication 112 may be implemented as a modified RRC Release message that includes one or more additional parameters configured to direct the WCD 18 to perform the desired functions. The wait-for-page communication 112 may also be implemented as another modified message in accordance with RRC protocol, as an entirely new RRC message, and/or as a message in accordance with another standard or protocol.

Meanwhile, as the WCD **18** idles in response to the wait-for-page communication **112**, the eNodeB **26** saves an indication of the service request included in message **110** and waits for the capacity of available network resources to increase. After the capacity of available network resources increases, the eNodeB **26** sends a page message **114** to the WCD **18** at time T_4 . Upon receiving the page message **110**, the WCD **18** sends a second PRACH preamble message **116**. The eNodeB then responds with a RAR **116**, and the WCD **18** sends another RRC Connection Request **120**. In response, the eNodeB **26** sends an RRC Connection Setup message **122**, at time T_5 .

The WCD **18** then transitions to connected mode, and sends an RRC Connection Setup Complete **124**, which is received by the eNodeB **26** at time T_6 . As before, the RRC Connection Setup Complete **124** can include a service request. So long as the eNodeB **26** continues to have sufficient network resources at time T_6 to accommodate the service request, the eNodeB **26** sends an S11 application message **126** to the MME **34** that includes the service request. In practice, the eNodeB **26** can undertake a similar determination as to whether to grant the second service request (in message **122**) as described above in connection with the initial service request (in message **110**). But because the second service request is sent in response to the eNodeB **26** determining that it has sufficient resources to accommodate the previously received service request (in message **110**), the eNodeB **26** is likely to grant the second service request unless, for example, there is a decrease in available network resources between times T_4 and T_6 or an increase in the quantity of network resources sought by the WCD **18**.

If the eNodeB **26** determines to grant the second service request (as in FIG. 2), the eNodeB **26**, MME **34**, and/or other network entities can then exchange additional control communications as necessary to establish communication links for bearer traffic and resources on the air interface. Such process may include, for example, authenticating the WCD **18**, establishing tunnel endpoint identifiers (TEIDs) to map packets of data between the WCD **18**, the eNodeB **26**, and the SGW **30**, configuring security parameters for such communications, as well as other functions. Once the communication link is established, the WCD **18** can exchange packets with remote entities as desired to thereby satisfy the service request.

III. Example Operations

FIGS. 3 and 4 are flow charts illustrating processes performed by separate network components to achieve efficient use of the access channel during loaded conditions by buffering incoming service requests. The process **50** shown in FIG. 3 may be performed by a base station, such as an eNodeB, according to an example embodiment. The process **70** shown in FIG. 4 may be implemented by another network component, such as a WCD, according to an example embodiment. It should be understood that the processes **50**, **70**, or portions thereof, may be implemented by other network components or combinations of network components, and/or may be implemented for other purposes, without departing from the scope of the present disclosure. For clarity, the various functions in the process **50** are described herein solely from the perspective of an eNodeB receiving a service request from a particular WCD, and the various functions in the process **70** are described herein solely from the perspective of a WCD seeking to connect with an LTE RAN via a particular eNodeB.

At block **52**, the eNodeB receives a service request from a WCD. The service request may be initiated by an unscheduled transmission on an access channel or otherwise involve

such an unscheduled transmission on the access channel. For example, the service request may be preceded by the random access procedure, which is initiated by a preamble message over PRACH. At block **54**, the eNodeB determines whether network channel resources are available to accommodate the service request. For example, the eNodeB may determine a quantity of RBs sought by the service request, and evaluate whether enough RBs are available during an upcoming interval to allocate to the WCD. As noted above in connection with FIG. 2, the quantity of network resources sought by the service request may be specified based on various factors, such as an amount of data to be carried, a quality of service class identifier (QCI), a bearer type, a minimum bit rate or latency, or another indicator related to a total number of RBs and/or a desired number of RBs per unit time. Accordingly, the eNodeB may make the determination at block **54** by evaluating the quantity of traffic channel RBs available in an immediate upcoming interval as well as on an ongoing basis and comparing that quantity with the quantity sought by the service request.

If, in block **54**, the eNodeB determines that network resources are available to accommodate the service request, then the process **50** proceeds to block **66**. At block **66**, the eNodeB allocates network resources to the WCD and provides a schedule of the allocated resources to the WCD. The eNodeB may also coordinate with other network entities to establish links within the RAN for carrying bearer traffic between the WCD and remote entities (e.g., tunnels for carrying communications between the WCD and a gateway system). Block **66** thus includes transitioning the WCD from idle mode to connected mode. Following block **66**, the WCD can engage in communications with remote entities using the allocated resources.

If, in block **54**, the eNodeB determines instead that resources are not available to accommodate the service request, then the process **50** proceeds to block **56**. At block **56**, the eNodeB saves an indication of the service request in a memory. For example, the eNodeB may save data indicative of the quantity of network resources sought by the service request and a unique identifier for the originating WCD to facilitate sending a page to that WCD subsequently. As used herein, saving data indicative of a received service request along with identifying information for the WCD that sent the service request is generally referred to as “buffering” the service request.

At block **58**, the eNodeB sends a wait-for-page communication to the WCD. As described above in connection with FIG. 2, the wait-for-page communication causes the WCD to forgo at least some transmissions on the access channel until the WCD receives a page message from the eNodeB. The prohibition on transmissions over the access channel may extend to access channel transmissions used to initiate a service request for the same or greater quantity of network resources than the quantity sought by the initial service request. Additionally or alternatively, the temporary prohibition caused by the wait-for-page communication may extend to all service requests, or to service requests seeking a quantity of network resources greater than a quantity specified by the wait-for-page communication. For instance, the eNodeB may include in the wait-for-page communication an indication of the quantity of its available network resources, and the WCD can then forgo service requests seeking a greater quantity.

At block **60**, the eNodeB waits for the capacity of its available network resources to increase. The wait block **60** may involve iteratively evaluating the capacity of available network resources (e.g., the quantity of available RBs) and,

for each iteration, determining whether the capacity is sufficient to accommodate the service request saved in memory. The wait block **60** may additionally or alternatively involve repetitively evaluating the capacity of available network resources, and comparing the capacity with a threshold value. Upon determining that the available capacity is sufficient to accommodate the saved service request (or exceeds a threshold), the process **50** may proceed with block **62**.

Meanwhile, as the eNodeB waits for its capacity of available network resources to increase, in block **60**, the WCD operates in idle mode and monitors a predetermined control channel for a page message. At block **62**, the eNodeB may send a page message to the WCD. The page message overrides the WCD's prohibition on certain access channel transmissions imposed by the wait-for-page communication, and thereby causes the WCD to send a second service request to the eNodeB. The second service request may be initiated by an unscheduled transmission on the access channel, such as a preamble message on PRACH.

At block **64**, the eNodeB receives the second service request from the WCD. In response to the second service request, the process **50** may return to block **54** to again determine whether network resources are available to accommodate the second service request. And then, upon determining that sufficient network resources are available, the process **50** may proceed with block **66** to setup the connection. At block **66** the eNodeB can allocate resources to the WCD so as to grant the second service request, establish links for carrying bearer traffic through the RAN, and provides the schedule of allocated network resources to the WCD, thereby transitioning the WCD to connected mode.

In some cases, the eNodeB may maintain a buffer of service requests from multiple WCDs, and can page each WCD in the order in which the original service requests were received. As a result the eNodeB can page the WCDs in order, which may help equalize latency associated with making initial connections among the WCDs served by the eNodeB. By contrast, in some systems, the eNodeB may make a connection with whichever WCD happens to submit a connection request immediately following an increase in available network resources. To track the order of reception of multiple service requests, the memory may store an indication of the reception time for each received service request, for example. Additionally or alternatively, each service request saved to memory can be associated with an index number of another incrementing value that can be used to determine the time order of each saved service request.

The WCDs associated with each service request in the buffer can then be paged in order once resources become available. For example, the eNodeB may evaluate each buffered service request individually and determine for each buffered service request whether available network resources are sufficient to accommodate the service request and, if not, wait until such resources become available. Upon determining that sufficient network resources are available to accommodate the next one of the service requests in the buffer, the eNodeB can then send a page message addressed to the WCD corresponding to that service request. Then, the eNodeB can repeat the process for the next buffered service request (e.g., the service request received next in time).

Referring now to FIG. 4, at block **72**, the WCD initiates connection with an eNodeB using an unscheduled transmission on an access channel. For example, the WCD may send a preamble message over PRACH to initiate the random access procedure described above. As described in connection with FIG. 2, in response to the unscheduled transmission, the WCD may receive a connection setup message from the

eNodeB, which resolves contention in the random access procedure, allocates initial uplink and downlink control channel resources, and transitions the WCD to connected mode. At block **74**, the WCD sends a service request for a particular quantity of network resources. As noted in FIG. 2, the service request may be included in a connection setup complete message (e.g., by including a buffers status report). The service request may specify the particular quantity of network resources sought by indicating a quality of service class identifier (QCI), bearer type, bit rate, latency, or another measure of a quantity and/or rate of resource blocks sought for the WCD.

At block **76**, the WCD receives a wait-for-page communication from the eNodeB in response to the service request. Upon receiving the wait-for-page communication, the WCD transitions to idle mode (e.g., adjusts the value of its state parameter), and monitors predetermined downlink control channels for a page message from the base station, at block **78**. While waiting for the page message (block **78**), the WCD forgoes further unscheduled transmissions to the eNodeB over the access channel for the purpose of initiating a service request for an equal or greater quantity of network resources than sought by the initial service request. Moreover, the WCD may forgo access channel transmissions for the purpose of initiating a service request for a quantity of network resources greater than a quantity indicated by the wait-for-page communication, or may simply forgo access channel transmissions entirely. The reduced access channel transmissions thereby reduce congestion over PRACH that would otherwise occur due to repeated attempts to complete the random access procedure.

The WCD interprets the wait-for-page communication as a directive to cause the WCD to operate as described in connection with block **78** (i.e., to forgo at least some use of the access channel). The wait-for-page communication may be implemented as an enhanced communication in accordance with the RRC protocol, such as an RRC Release or RRC Rejection message with one or more extra parameters. Upon receipt, the WCD may implement the directive by temporarily removing the particular cell and/or particular base station from a listing of available cells the WCD is allowed to connect to. For example, the WCD may maintain a listing of neighboring cells, perhaps derived in part from system information broadcasts, and also maintain access control parameters associated with each cell. The access control parameters can specify whether or not the WCD is permitted to initiate connection with those cells (e.g., by initiating the random access procedure and sending a connection request). Implementing the directive from the wait-for-page communication may involve adjusting the access control parameters (until receipt of a subsequent page) such that the WCD is barred from sending another connection request to the same cell for the same connection type, or for a connection type that demands an even greater quantity of network resources.

In addition to waiting for the page message from the particular eNodeB, in block **78**, the WCD may also attempt to initiate a connection with another base station. For example, after receiving the wait-for-page communication, the WCD may monitor pilot signals from another eNodeB with an overlapping coverage area, and attempt to register with such other eNodeB. The WCD may send a registration request and/or access request to the other eNodeB. Of course, in practice, registering with the other eNodeB generally involves using the random access procedure to initiate communications with the other eNodeB. To allow the WCD to proceed as such, the wait-for-page communication's instruction to forgo further access channel transmissions can be

implemented with eNodeB specificity. For example, the WCD may forgo sending service requests (preceded by access channel transmissions) to the particular eNodeB that sent the wait-for-page communication, but still send access channel transmissions to other eNodeBs.

In the event the WCD successfully forms a connection with another eNodeB during block 78, the process 70 may end and the WCD may ignore a subsequent page message from the particular eNodeB. Additionally or alternatively, upon the WCD being registered with the network via another eNodeB, the MME (or another network entity) may notify the particular eNodeB and that eNodeB can then remove the pending service request from its memory. Such communications may occur, for example, in response to the WCD informing the LTE RAN (after connecting through the other eNodeB) that the WCD previously received a wait-for-page communication from the particular eNodeB.

However, in circumstances in which the WCD is unsuccessful in connecting to another base station, or in which there are no base stations with overlapping coverage areas, the WCD continues to wait in idle mode until it receives a page. Then, at block 80, the WCD receives a page message from the eNodeB. The page message allows the WCD to again utilize the access channel without restriction. In other words, the overrides the earlier instruction imposed by the wait-for-page communication to forgo (at least some) transmissions using the access channel. In response to receiving the page message, at block 80, the WCD again initiates connection with the eNodeB using an unscheduled transmission on the access channel, at block 82. Block 82 may involve, for example, the WCD sending a preamble message over PRACH. At block 84, the WCD sends a second service request for network resources. Similar to the first service request, the second service request may include a buffer status report that indicates data ready to be sent by the WCD and may indicate a quality of service class identifier, a minimum bit-rate, latency, or the like to specify a desired quantity and/or rate of RBs sought by the WCD.

The eNodeB can then determine to grant the second service request, and then exchange signals with the MME and/or other network entities to establish links for bearer communications and allocate resources for communicating with the WCD over the eNodeB's air interface. At block 86, the WCD receives a schedule of the allocated resources from the eNodeB and engages in communication over the LTE RAN using those resources.

IV. Example RAN Components

FIG. 5 is a simplified block diagram of an example RAN component 91, according to an example embodiment. In particular, FIG. 5 illustrates functional components that might be found in a RAN component 91 that is arranged to operate in accordance with the embodiments described above. As shown, the RAN component 91 may include a communication interface 90, a processing unit 92, and data storage 94, all of which may be communicatively linked together by a system bus, network, or one or more other connection mechanisms 96. The data storage 94 may include a non-transitory computer readable medium and includes program logic 98 that, when executed by the processing unit 92, cause the RAN component 91 to function in accordance with the present disclosure.

In practice, RAN component 91 may take the form of an eNodeB or a WCD, or may take the form of another component of an LTE network. Further, the illustrated components of RAN component 91 (e.g., communication interface 90, a processing unit 92, and/or data storage 94) may be distributed and/or subdivided between one or more entities in an LTE

network and/or in another network. It should be understood that an example system may also take the form of another network entity or combinations of other network entities, without departing from the scope of the present disclosure.

In RAN component 91, communication interface 90 may comprise one or more or wired or wireless communication interfaces and/or other associated equipment for engaging in communications with other network entities and/or for engaging in RF communications with mobile stations according to one or more air interface protocols. Thus, the network interface 90 may include antenna structures configured to send and receive radio communications to define an interface (e.g., the LTE air interface 36 described above in connection with FIG. 1B). The communication interface 90 may comprise any sort of communication link or mechanism enabling the RAN component 91 to exchange signaling and bearer data with other network entities, such as used to create tunnels for passing IP packets between the PGW 32 and WCD 18. Further, processing unit 92 may comprise one or more processors (e.g., general purpose and/or special purpose processors), such as microprocessors for instance.

Data storage 94 may be a non-transitory computer readable medium. For example, data storage 94 may take the form of one or more volatile and/or non-volatile storage components, such as magnetic, optical, or organic storage components, integrated in whole or in part with processing unit 92. As further shown, data storage 94 contains program logic 98 (e.g., machine language instructions) executable by processing unit 92 to carry out various functions, such as the functionality of the example methods and systems described herein.

For instance, in an example in which the networking node 91 is implemented as a base station, the communication interface 90 may include one or more antenna structures configured to wirelessly communicate with WCDs served by the base station over one or more scheduled channels and over an unscheduled access channel. The base station can be configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs. A given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station can be configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel. Also, the processing unit 92 may be operatively coupled to the antenna structures of the base station and the processing unit 92 and the program logic 98 may form a controller configured to (i) receive, at the base station, a first service request from a particular WCD, wherein the first service request specifies a particular quantity of network resources; (ii) responsive to receiving the first connection request, make a first determination that a capacity of available network resources on the scheduled channels is insufficient to allocate the particular quantity of network resources to the particular WCD; and (iii) responsive to making the first determination: (a) send a wait-for-page communication to the particular WCD, wherein the wait-for-page communication causes the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the particular quantity of network resources; (b) store data indicative of the received service request in a memory; (c) wait for the capacity of available network resources on the scheduled channels to increase; (d) after waiting for the capacity of available net-

23

work resources on the scheduled channels to increase, make a second determination that the capacity of available network resources on the scheduled channels is sufficient to allocate the particular quantity of network resources to the particular WCD; and (e) responsive to making the second determination, send a page to the particular WCD. 5

Example embodiments have been described above. It should be understood, however, that variations from these embodiments are possible, while remaining within the true spirit and scope of the disclosure. 10

We claim:

1. A method of operating a wireless network system, wherein the wireless network system includes a base station operating to serve wireless communication devices (WCDs) via one or more scheduled channels and an unscheduled access channel, wherein the base station is configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs, the method comprising:
 - receiving a first service request from a particular WCD, wherein the first service request specifies a particular quantity of network resources; 25
 - responsive to receiving the first service request, making a first determination that a capacity of available network resources on the scheduled channels is insufficient to allocate the particular quantity of network resources to the particular WCD; and 30
 - responsive to making the first determination, the base station:
 - (i) sending a wait-for-page communication to the particular WCD, wherein the wait-for-page communication causes the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the particular quantity of network resources, wherein prior to sending the wait-for-page communication, determining a quantity of available network resources on the scheduled channels and wherein the wait-for-page communication includes an indication of the determined quantity of available network resources so as to further cause the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the determined quantity of available network resources; 40
 - (ii) storing data indicative of the received service request in a memory; 45
 - (iii) waiting for the capacity of available network resources on the scheduled channels to increase; and 50
 - (iv) responsive to waiting, sending a page to the particular WCD. 55
2. The method of claim 1, further comprising:
 - responsive to sending the page to the particular WCD, receiving a second service request from the particular WCD, wherein the second service request is initiated by an unscheduled transmission over the access channel from the particular WCD to the base station; and
 - responsive to receiving the second service request, allocating network resources on the scheduled channels for the particular WCD and sending the particular WCD a schedule of the allocated network resources. 65

24

3. The method of claim 2, wherein the base station schedules the available network resources of the scheduled channels in resource blocks that each span a non-overlapping time-frequency segment of an available spectrum on the scheduled channels; and

wherein allocating network resources on the scheduled channels for the particular WCD comprises: (i) determining a quantity of resource blocks sufficient to accommodate the second service request, and (ii) reserving the determined quantity of resource blocks for the particular WCD.

4. The method of claim 1, wherein the wait-for-page communication further causes the particular WCD to forgo transmitting to the base station over the access channel prior to the particular WCD receiving the page message.

5. The method of claim 1, and wherein a given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station is configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel, and

wherein each of the WCDs served by the base station is configured to operate alternately: in a connected mode, in which the WCD communicates with the base station using network resources, on the scheduled channels, that are allocated by the base station, and in an idle mode, in which the WCD monitors predetermined channels for pages and does not have allocated network resources on the scheduled channels, the method further comprising: prior to receiving the first service request:

- (i) receiving, via the access channel, an unscheduled transmission from the particular WCD;
- (ii) sending the particular WCD a connection setup communication that causes the particular WCD to transition to connected mode; and

wherein the wait-for-page communication further causes the particular WCD to transition to idle mode.

6. The method of claim 1, further comprising: comparing the capacity of available network resources on the scheduled channels with the particular quantity of network resources specified by the first service request; and

basing the first determination, at least in part, on the comparison of the capacity of available network resources and the quantity of network resources specified by the first service request.

7. The method of claim 1, wherein waiting for the capacity of available network resources on the scheduled channels to increase comprises: at least once, (i) monitoring the capacity of available network resources on the scheduled channels, (ii) comparing the monitored capacity to the particular quantity of network resources specified by the first service request, and (iii) determining, based on the comparison, whether the monitored capacity exceeds the particular quantity of network resources; and

wherein sending the page message is performed responsive to determining that the monitored capacity exceeds the particular quantity of network resources.

8. The method of claim 1, further comprising: the base station receiving a plurality of additional service requests, each at a respective reception time from a respective WCD, wherein each of the additional service requests specifies a respective quantity of network resources;

25

sending respective wait-for-page communications to the respective WCDs so as to cause each of the WCDs to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the respective quantity of network resources;

storing an indication of the additional service requests in the memory such that each of the additional service requests is associated with its order of reception at the base station; and

sending pages to the respective WCDs in the order in which their associated service requests are received at the base station, such that each given page is sent to a given WCD responsive to determining that the capacity of available network resources on the scheduled channels is sufficient to accommodate the service request from the given WCD.

9. A method of operating a wireless communication device (WCD) in a wireless network system, wherein the wireless network system includes a base station operating to serve a plurality of WCDs including the WCD via one or more scheduled channels and an unscheduled access channel, wherein the base station is configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs, the method comprising:

sending a first service request from the WCD to the base station, wherein the first service request specifies a particular quantity of network resources;

responsive to sending the first service request, receiving a wait-for-page communication from the base station;

responsive to receiving the wait-for-page communication, the WCD, prior to being paged, forgoing transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the particular quantity of network resources;

receiving a page from the base station; and

responsive to receiving the page, sending a second service request to the base station, wherein the second service request is initiated by the WCD sending an unscheduled transmission over the access channel to the base station.

10. The method of claim **9**, further comprising, responsive to receiving the wait-for-page communication, and prior to receiving the page, the WCD further forgoing all transmitting to the base station over the access channel.

11. The method of claim **9**, wherein the wait-for-page communication includes an indication of a quantity of available network resources, the method further comprising:

responsive to receiving the wait-for-page communication, and prior to receiving the page, the WCD further forgoing transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the indicated quantity of available network resources.

12. The method of claim **9**,

wherein a given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station is configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel; and

wherein the WCD is configured to operate alternately: in a connected mode, in which the WCD communicates with

26

the base station using network resources, on the scheduled channels, that are allocated by the base station, and in an idle mode, in which the WCD monitors predetermined channels for page messages and does not have allocated network resources on the scheduled channels, the method further comprising:

prior to sending the first service request, the WCD:

- (i) while in idle mode, sending, via the access channel, an unscheduled transmission to the base station;
 - (ii) responsive to sending the unscheduled transmission, receiving a connection setup communication from the base station; and
 - (iii) responsive to receiving the connection setup message, transitioning from idle mode to connected mode; and
- responsive to receiving the wait-for-page communication, transitioning to idle mode.

13. The method of claim **9**, further comprising: responsive to sending the second service request, receiving a schedule of allocated network resources.

14. A wireless network system comprising:

a base station having one or more antenna structures configured to wirelessly communicate with wireless communication devices (WCDs) served by the base station over one or more scheduled channels and over an unscheduled access channel, wherein the base station is configured to allocate available network resources amongst the WCDs by scheduling network traffic over the scheduled channels so as to accommodate respective demands of the WCDs indicated, at least in part, by respective service requests received from the WCDs, and wherein a given WCD that does not have allocated network resources on the scheduled channels with which to communicate to the base station is configured to initiate a connection with the base station by sending an unscheduled transmission to the base station over the access channel; and

a controller operatively coupled to the antenna structures of the base station, wherein the controller is configured to:

- (i) receive, at the base station, a first service request from a particular WCD, wherein the first service request specifies a particular quantity of network resources;
- (ii) responsive to receiving the first connection request, make a first determination that a capacity of available network resources on the scheduled channels is insufficient to allocate the particular quantity of network resources to the particular WCD; and

(iii) responsive to making the first determination:

- (a) send a wait-for-page communication to the particular WCD, wherein the wait-for-page communication causes the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the particular quantity of network resources, wherein prior to sending the wait-for-page communication, determine a quantity of available network resources on the scheduled channels, and include, in the wait-for-page communication, an indication of determined quantity of available network resources so as to further cause the particular WCD to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the determined quantity of available network resources;
- (b) store data indicative of the received service request in a memory;

27

- (c) wait for the capacity of available network resources on the scheduled channels to increase; and
- (d) responsive to waiting, send a page to the particular WCD.

15. The wireless network system of claim 14, wherein the controller is further configured to:

responsive to sending the page to the particular WCD, receive a second service request from the particular WCD, wherein the second service request is initiated by an unscheduled transmission over the access channel from the particular WCD to the base station; and responsive to receiving the second service request, allocate network resources on the scheduled channels for the particular WCD and send the particular WCD a schedule of the allocated network resources.

16. The wireless network system of claim 15, wherein the controller is further configured to:

schedule the available network resources of the scheduled channels in resource blocks that each span a non-overlapping time-frequency segment of an available spectrum on the scheduled channels; and

allocate network resources on the scheduled channels for the particular WCD by: (i) determining a quantity of resource blocks sufficient to accommodate the second service request, and (ii) reserving the determined quantity of resource blocks for the particular WCD.

17. The wireless network system of claim 14,

wherein each of the WCDs served by the base station is configured to operate alternately: in a connected mode, in which the WCD communicates with the base station using network resources, on the scheduled channels, that are allocated by the base station, and in an idle mode, in which the WCD monitors predetermined channels for pages and does not have allocated network resources on the scheduled channels,

28

wherein the controller is further configured to, prior to receiving the first service request:

(i) receive, via the access channel, an unscheduled transmission from the particular WCD; and

(ii) send the particular WCD a connection setup communication that causes the particular WCD to transition to connected mode; and

wherein the wait-for-page communication further causes the particular WCD to transition to idle mode.

18. The wireless network system of claim 14, wherein the controller is further configured to:

(i) receive, at the base station, a plurality of additional service requests, each at a respective reception time from a respective WCD, wherein each of the additional service requests specifies a respective quantity of network resources;

(ii) send respective wait-for-page communications to the respective WCDs so as to cause each of the WCDs to, prior to being paged, forgo transmitting to the base station over the access channel for the purpose of initiating a service request specifying a greater quantity of network resources than the respective quantity of network resources;

(iii) store an indication of the additional service requests in the memory such that each of the additional service requests is associated with its order of reception at the base station; and

(iv) send pages to the respective WCDs in the order in which their associated service requests are received at the base station, such that each given page is sent to a given WCD responsive to determining that the capacity of available network resources on the scheduled channels is sufficient to accommodate the service request from the given WCD.

* * * * *