

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
15 January 2004 (15.01.2004)

PCT

(10) International Publication Number
WO 2004/006537 A2

(51) International Patent Classification⁷: **H04L 29/06**

(21) International Application Number:
PCT/US2003/021116

(22) International Filing Date: 8 July 2003 (08.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/192,919 10 July 2002 (10.07.2002) US

(71) Applicant: **CISCO TECHNOLOGY, INC.** [US/US];
170 West Tasman Drive, San Jose, CA 95134-1706 (US).

(72) Inventors: **O'ROURKE, Chris**; 602 Scotts Ridge Trail, Apex, NC 27502 (US). **BATZ, Robert, M.**; 5508 Harrington Grove Drive, Raleigh, NC 27613 (US). **DABBOUSSI, Rabilh, A.**; 107 Kindred Way, Cary, NC 27513 (US). **GLOTZER, John, M.**; 202 Cates Farm Road, Chapel Hill, NC 27516 (US). **MENDITTO, Louis, F.**; 4701 Tanglewood Drive, Raleigh, NC 27612 (US). **PATEL, Alpesh, S.**; 1901 Halford Ave # 184, Santa Clara, CA 95051 (US). **LEUNG, Kent, K.**; 2447 Villa Nueva Way, Mountain View, CA 94040 (US).

(74) Agent: **SHOWALTER, Barton, E.**; Baker Botts, L.L.P., 2001 Ross Avenue, Dallas, TX 75201-2980 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK (utility model), SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for all designations
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations

Published:

- without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR COMMUNICATING IN A LOADBALANCING ENVIRONMENT

(57) Abstract: A method for communicating in a loadbalancing environment is provided that in a particular embodiment includes receiving a request packet from a network access server (NAS) to initiate a communication session. The request packet is then communicated to a tunneling protocol network server (TPNS) and a response packet is received in response to the request packet. The response packet establishes a tunnel that facilitates the communication session and that includes an identification element associated with the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS.



WO 2004/006537 A2

SYSTEM AND METHOD FOR COMMUNICATING
IN A LOADBALANCING ENVIRONMENT

TECHNICAL FIELD OF THE INVENTION

This invention relates in general to communications and more particularly to system and method for communicating in a loadbalancing environment.

BACKGROUND OF THE INVENTION

Networking architectures are growing increasingly complex in communications environments. In addition, the augmentation of clients or end users wishing to communicate in a network environment have caused many networking architectures and systems to respond by adding elements to accommodate the increase in network traffic. Communication tunnels may be used in order to establish or to gain access to a network, whereby an end user or a client may initiate the corresponding tunneling protocol by invoking a central location or a single network node. Such central locations may be obligated to execute an excessive number of tasks that overburden or overtax the central location.

In addition, the unreasonable delegation of a disproportionate number of duties may decrease throughput, thereby inhibiting the flow of network traffic because of the robust communications propagating through the network. These strains on the central location not only operate to slow network traffic, but

such reliance on a single central node in a network may create severe problems in situations where the central node fails or is otherwise unable to perform all of its assigned duties. In egregious cases, network
5 communications associated with the central location may be lost and unrecoverable when the central location is overwhelmed with data.

SUMMARY OF THE INVENTION

10 From the foregoing, it may be appreciated by those skilled in the art that a need has arisen for an improved communications approach that provides for a reduction in the burden on a loadbalancer associated with communications between two endpoints. In accordance with
15 one embodiment of the present invention, a system and method for establishing a network communication in a loadbalancing environment are provided that substantially eliminate or greatly reduce disadvantages and problems associated with conventional loadbalancing techniques.

20 According to one embodiment of the present invention, there is provided a method for communicating in a loadbalancing environment that includes receiving a request packet from a network access server (NAS) to initiate a communication session. The request packet is
25 then communicated to a tunneling protocol network server (TPNS) and a response packet is received in response to the request packet. The response packet establishes a tunnel that facilitates the communication session and that includes an identification element associated with
30 the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS.

Certain embodiments of the present invention may provide a number of technical advantages. For example, according to one embodiment of the present invention a communications approach is provided that allows a loadbalancer to only be actively involved in the initiation of a communication session. This reduction in responsibility for the loadbalancer operates to increase throughput as two points or nodes communicate more directly instead of having to direct all information through the loadbalancer for processing. This may further reduce the number of central processing unit (CPU) cycles, which may be intensive and require additional work to be performed by the loadbalancer. Accordingly, the loadbalancer may be relegated to simple information transfer or packet switching (generally not involving a modification of address information). The removal of the loadbalancer from the data transfer interaction between two nodes may further alleviate responsibilities designated for the loadbalancer such that more robust network traffic may be accommodated in the network.

Yet another technical advantage of one embodiment of the present invention is also a result of the removal of the loadbalancer from data transfer traffic. The decreased reliance on the loadbalancer operates to better allocate memory resources because information per-tunnel state is no longer maintained by the loadbalancer. In addition, the decreased dependency on the loadbalancer allows for improved failover characteristics whereby, if the loadbalancer would become dysfunctional or non-operational, associated communications sessions would not be lost. The continued operation of existing

communication tunnels may be preserved in accordance with the teachings of the present invention by relegating the loadbalancer to tasks involving only the initiation of the communications session. Embodiments of the present invention may enjoy some, all, or none of these advantages. Other technical advantages may be readily apparent to one skilled in the art from the following figures, description, and claims.

10 BRIEF DESCRIPTION OF THE DRAWINGS

To provide a more complete understanding of the present invention and features and advantages thereof, reference is made to the following description, taken in conjunction with the accompanying figures, wherein like reference numerals represent like parts, in which:

FIGURE 1 is a simplified block diagram of a system for communicating in a loadbalancing environment in accordance with one embodiment of the present invention;

FIGURE 2 is a simplified state diagram of the operation of the system for communicating in a loadbalancing environment;

FIGURE 3 is a simplified block diagram of a system for communicating in a loadbalancing environment in accordance with another embodiment of the present invention; and

FIGURE 4 is a flowchart illustrating a series of example steps associated with a method for communicating in a loadbalancing environment.

DETAILED DESCRIPTION OF THE INVENTION

FIGURE 1 is a simplified block diagram illustrating a communication system 10 for communicating data in a loadbalancing environment. Communication system 10 includes an end user 12, a radio access network (RAN) 14, a network access server (NAS) 18, and an access network 22. Additionally, communication system 10 includes a loadbalancer 26, multiple tunneling protocol network servers (TPNSs) 30a-n, an authentication, authorization, and accounting (AAA) server 32, an internet protocol (IP) network 36, and a destination server 38.

In accordance with the teachings of the present invention, communication system 10 operates to alleviate the responsibilities associated with loadbalancer 26 in providing optimal communications between end user 12 and IP network 36. Two stages generally exist in communications flows that involve end user 12. A first stage relates generally to initiation whereby a communication session may be prompted by end user 12. A second stage relates generally to the establishment of the communication session or link with corresponding data transfer. The initiation stage of a communication session generally requires an invocation of loadbalancer 26. During this stage, a create request and a suitable response may be generated within communication system 10.

After the communication session is initiated, loadbalancer 26 may be removed from the communications pathway allowing a more direct data transfer between end user 12 and IP network 36 via access network 22. Loadbalancer 26 operates to be only involved in initiating of the communication session in accordance with the teachings of the present invention whereby

loadbalancer 26 operates as a request broker for establishing communication tunnels. This reduces the number of required processing cycles and additionally preserves memory resources of loadbalancer 26 since it no longer needs to maintain per-tunnel state information. This implementation also provides for a better allocation of resources associated with loadbalancer 26.

Additionally, when this initiation and establishment protocol is implemented in communication system 10, loadbalancer 26 experiences a significant reduction in communications traffic. This in turn alleviates the burden placed on loadbalancer 26 and offers increased throughput and decreased reliance on a central location (i.e. loadbalancer 26). This feature further provides the opportunity for loadbalancer 26 to be isolated from extensive central processing unit (CPU) cycles, that may be intensive and require substantial work to be performed by loadbalancer 26. This technique also allows loadbalancer 26 to participate in only simple information transfers or data packet switching without requiring loadbalancer 26 to modify addressing information or to process data passing therethrough. With the reduced responsibilities of loadbalancer 26, communication system 10 is able to accommodate more robust communications traffic and allow for an enhanced and more efficient communications environment.

End user 12 is a client or a customer wishing to initiate a communication in communication system 10 via access network 22. End user 12 may be inclusive of devices used to initiate a communication, such as a computer, a personal digital assistant (PDA), a laptop or electronic notebook, a telephone, or any other device,

component, element, or object capable of initiating voice or data exchanges within communication system 10. End user 12 may also be inclusive of a suitable interface to the human user, such as a microphone, a display, or a keyboard or other terminal equipment (such as for example an interface to a personal computer or to a facsimile machine in cases where end user 12 is used as a modem). End user 12 may also be any device that seeks to initiate a communication on behalf of another entity or element, such as a program, a database, or any other component, device, element, or object capable of initiating a voice or a data exchange within communication system 10. Data, as used herein in this document, refers to any type of numeric, voice, or script data, or any type of source or object code, or any other suitable information in any appropriate format that may be communicated from one point to another.

RAN 14 is a communications interface between end user 12 and NAS 18. RAN 14 may comprise a base transceiver station and a base station. The communications interface provided by RAN 14 allows data to be exchanged between end user 12 and any number of selected elements within communication system 10. RAN 14 facilitates the delivery of a request packet generated by end user 12 and the reception of information sought by end user 12. RAN 14 is only one example of a communications interface between end user 12 and NAS 18. Other types of communications interfaces may be used for a desired network design.

NAS 18 is an element that provides access to any network (such as access network 22) for end user 12. NAS 18 may be used with a transmission control

protocol/internet protocol (TCP/IP) network, including serial terminal access controllers, modem pools or stacks, integrated services digital network (ISDN) routers, or multi-function access controllers where appropriate. NAS 18 may also be used in combination with any element that provides switched service connections, point-to-point (PPP) serial IP protocols, or user authentication functions according to particular needs. NAS 18 may represent a tunnel initiator for communication system 10, whereby TPNSs 30a-n represent multiple tunnel terminators or contact nodes for establishing a communication session. NAS 18 may also support serial line internet protocol (SLIP) and allow NAS 18 to establish and to manage the individual communications links to remote sites across a switched service. NAS 18 may properly authenticate end user 12 by invoking AAA server 32 before allowing access to a network or to another server. NAS 18 may also store one or more identification elements or passwords that may be used in authenticating end user 12.

In a particular embodiment of the present invention, the communication protocol implemented by NAS 18 is RADIUS. NAS 18 may alternatively use terminal access controller access control system (TACACS), or diameter, or any other suitable communications protocol in order to provide an authentication functionality to end user 12. In operation, NAS 18 operates to bring up a communication session with end user 12. NAS 18 may also provide accounting or authorization functions on behalf of end user 12 and perform IP address management for end user 12 where appropriate. NAS 18 may terminate PPP connections or communication links and may generally correspond to

the communication protocol implemented by access network 22.

In an alternative embodiment of the present invention, NAS 18 may be removed from communication system 10 or substituted with any element capable of performing one or more of the functions of NAS 18. NAS 18 could be removed in various applications such as for example with use in a cable implementation via PPP over Ethernet (PPPOE) where data may be tunneled over or across the PPPOE connection. The data may be communicated to an element like NAS 18 that provides an interface between access network 22 and end user 12. Alternatively, IP packets (not tunneled in PPPOE) may be routed from the time they arrive on the network and directly sent to an element that receives the data. Additionally, some or all of the functions of NAS 18 may be provided in access network 22. Accordingly, access network 22 may be configured in any appropriate manner in order to provide PPP or RADIUS-type mechanisms or features, such as authentication, authorization, accounting, content filtering, and priority for example.

Access network 22 represents a series of points or nodes of interconnected communication paths for receiving and transmitting packets of information that propagate through communication system 10. Access network 22 offers a communicative interface between end user 12 and IP network 36 and may provide a PPP connection in certain embodiments. Access network 22 may implement any communications protocol such as dial, cable, digital subscriber line (DSL), fiberoptic, radio, local area network (LAN), wireless local area network (WLAN), metropolitan area network (MAN), wide area network (WAN),

or any other suitable communications architecture or platform that allows packet communications or tunneling to (or through) access network 22. Access network 22 may also be inclusive of RAN 14 where appropriate or a hub
5 that allows end user 12 to log onto or otherwise access a network.

Access network 22 may also include authentication features provided to end user 12. In a particular embodiment, access network 22 represents a packet data
10 network (PDN). However, access network 22 may be any other suitable network where appropriate and according to particular needs. Access network 22 implements a TCP/IP communications language architecture in a particular embodiment of the present invention. However, access
15 network 22 may alternatively implement any other suitable communications protocol for transmitting and receiving data packets within communication system 10.

Loadbalancer 26 is an element or a device that receives requests and then distributes those requests to
20 the next available server or node. The available server or node may be any computer or device on a network that manages network resources or that processes data. Loadbalancer 26 performs loadbalancing, which refers to the dividing of the amount of work that an element has to
25 do between two or more elements such that more work gets done in the same amount of time and, in general, end users are served more quickly. Loadbalancer 26 may be implemented with hardware, software, (or a combination of both) or any component, device, element, or object that
30 suitably manages information traffic in a network environment. Any of the operations of NAS 18 or TPNSSs 30a-n (or the actual elements themselves) may be provided

in loadbalancer 26 where appropriate and in accordance with particular needs.

TPNSSs 30a-n each represent a server program or element that may be provided to address secure services provided to end user 12. TPNSSs 30a-n may each provide an interface between IP network 36 and loadbalancer 26. TPNSSs 30a-n may also provide a more direct connection to access network 22 (illustrated as a set of dashed lines 24 in FIGURE 1) such that the responsibilities relegated to loadbalancer 26 are significantly reduced. TPNSSs 30a-n may represent a contact node or tunnel terminator provided to communication system 10. In a particular embodiment, TPNSSs 30a-n offer secure services within communication system 10. In such a case, a corresponding local access concentrator (LAC) element may be provided in any one or more of the elements within communications system 10. Alternatively, TPNSSs 30a-n may facilitate non-secure services or communications involving end user 12 where appropriate and according to particular needs. It is critical to note that TPNSSs 30a-n may be any suitable component, hardware, software, object, or element that offers a communications node for end user 12 to initiate contact therewith. TPNSSs 30a-n represent points of great flexibility in that they may be any suitable element that facilitates communications via loadbalancer 26.

AAA server 32 is a server program that handles requests by end user 12 for access to networking resources. Networking resources refers to any device, component, or element that provides some operation or functionality to end user 12 communicating in communication system 10. For a corresponding network,

AAA server 32 may also provide authentication, authorization, and accounting services and management. Authorization generally refers to the process of giving end user 12 permission to do or to access something. In multi-user computer systems, a system administrator may define for the system which end users are allowed access to given data in the system and, further, what privileges are provided for end user 12. Once end user 12 has logged into a network, such as for example IP network 36 or access network 22, the network may wish to identify what resources end user 12 is given during the communication session. Thus, authorization within communication system 10 may be seen as both a preliminary setting of permissions by a system administrator and the actual checking or verification of the permission values that have been set, when end user 12 attempts access. Authentication generally refers to the process of determining whether end user 12 is in fact who or what it is declared to be. In the case of private or public computer networks, authentication may be commonly done through the use of unique identification elements (such as an MSISDN) or log-on passwords. Knowledge of the password offers a presumption that a given end user is authentic. Accounting generally refers to tracking usage for each end user or each network and may additionally include trafficking information or data relating to other information flows within communication system 10 or within a particular sub-network.

AAA server 32 may receive the IP address and other parameters from any suitable source, such as a client aware element or alternatively from a dynamic host configuration protocol (DHCP) server or a domain name

system (DNS) database element, in order to direct data to be communicated to end user 12. AAA server 32 may include any suitable hardware, software, component, or element that operates to receive data associated with end user 12 and provides corresponding AAA related functions to network components within communication system 10. Authorization and IP address management may be retrieved by AAA server 32 from a selected TPNSs 30a-n where appropriate and may be provided to address secure services for end user 12. The assigned IP address may be a private or a routable IP address. On assignment of the IP address, the DHCP server may perform update procedures for updating the assigned IP address and leasing parameters for end user 12 in accordance with particular needs of the network.

IP network 36 represents a series of points or nodes of interconnected communication paths for receiving and transmitting packets of information that propagate through communication system 10. IP network 36 offers a communicative interface between destination server 38 and a selected TPNS 30a-n and may be any LAN, WLAN, MAN, WAN, or any other appropriate architecture or system that facilitates communications in a network environment. IP network 36 implements a TCP/IP communication language protocol in a particular embodiment of the present invention. However IP network 36 may alternatively implement any other suitable communication protocol for transmitting and receiving data packets within communication system 10.

Destination server 38 represents a program that, using the client/server model and the world wide web's HTTP, serves the files that form web pages to web users.

Destination server 38 may be provided as part of a larger package of internet and intranet-related programs for serving e-mail, downloading requests for file transfer protocol (FTP) files, building and publishing web pages, or any other suitable network operations according to particular needs. Alternatively, destination server 38 may be any suitable database or location in a network environment sought to be contacted, queried, or otherwise accessed by end user 12.

FIGURE 2 is a simplified state diagram of the operation of communication system 10 in accordance with one embodiment of the present invention. In operation, end user 12 initiates a connection to NAS 18 (possibly via RAN 14) which may include the use of PPP sessions and providing network access to end user 12. The PPP link could be a public switched telephone network (PSTN) connect, a provisioned asynchronous transfer mode (ATM) permanent virtual circuit (PVC) connection (or a switched virtual circuit (SVC), wherein the identification element includes an internet protocol (IP) address associated with the corresponding TPNS such that the NAS may communicate the data transfer to the TPNS connection), segments of the communication spectrum used for data channels obtained over a mobile wireless RAN, or any other suitable communications link according to particular needs. In response to this initiation, NAS 18 establishes a tunnel to a selected TPNS 30a-n via access network 22 and through loadbalancer 26. This is illustrated by steps 1 and 2 of FIGURE 2. This tunnel establishment phase allows a selected TPNS 30a-n to authenticate and authorize end user 12 in a central location without burdening multiple NAS 18 structures.

After end user 12 is authorized by a selected TPNS 30a-n, access may be granted to IP network 36 by the selected TPNS 30a-n to end user 12. The selected TPNS 30a-n may access AAA server 32 in granting permissions to communicate with IP network 36. In accordance with the teachings of the present invention, loadbalancer 26 may be only involved in tunnel establishment and generally not implicated in the subsequent data transfer. Communications protocols may be provided with the ability to assign or designate data transfer locations in the response signal generated during tunnel establishment illustrated by steps 3 and 4 of FIGURE 2. An identification element (such as an IP address, a data segment associated with the location of the selected TPNS 30a-n, or any other suitable element that operates to distinguish information routing or information propagation in a network) may be inserted into the response packet such that more direct communications may take place between NAS 18 and the selected TPNS 30a-n. Accordingly, for such protocols, loadbalancer 26 is only involved in the tunneling setup phase because NAS 18 will send data transfer messages directly to the selected TPNS 30a-n and bypass loadbalancer 26 as illustrated by step 5 of FIGURE 2. Using this method of loadbalancing for these tunnels, loadbalancer 26 may effectively step out of the way of data flows on the tunnel. This provides for improved scalability by limiting the packet inspection or modification duties of loadbalancer 26.

FIGURE 3 is a simplified block diagram of a system for communicating in a loadbalancing environment in accordance with another example embodiment of the present invention. FIGURE 3 illustrates an alternative

communications platform for mobile IP communications in which multiple foreign agents 54a-n are included in the system, as well as multiple home agents 60a-n. These elements cooperate in order to reduce the burden on
5 loadbalancer 26 such that loadbalancer 26 is only involved in the initiation of a communications session involving end user 12.

Home agents 60a-n each represent a mobile IP element or node that allows end user 12 to communicate in a
10 network environment. Each home agent 60a-n provides an interface between loadbalancer 26 and IP network 36 and communicates with loadbalancer 26 primarily during initiation of a communication session involving end user 12. Home agents 60a-n may be provisioned in any suitable
15 location of a network. Home agents 60a-n may cooperate with a DHCP server during mobile IP registration in order to assign an IP address to end user 12. User authentication and IP address allocation may be performed during mobile IP registration with a selected home agent
20 60a-n (this is slightly different from the PPP establishment phase for a simple IP service scenario). On authentication, end user 12 may be assigned an IP address by the selected home agent 60a-n and a corresponding network registrar where appropriate. The
25 assigned IP address may be a private or a routable IP address in accordance with particular needs.

In operation, where mobile IP services are being offered to end user 12, a home network may perform user authentication and IP address allocation. User
30 authentication and IP address allocation may be performed during mobile IP registration with the selected home agent 60a-n. Mobile IP enables a host to be identified

by a single IP address even while end user 12 physically moves its point of attachment from one network (or point in the network) to another. This feature allows transparent forwarding of data packets to end user 12.

5 Movement from one point of attachment to another is seamlessly achieved without requiring the intervention of end user 12. Thus, mobile IP servicing in the context of communication system 10 provides ubiquitous connectivity for users irrespective of their presence in their

10 respective home enterprise networks.

Foreign agents 54a-n are routing elements or entities that facilitate communications initiated by end user 12. Elements or devices associated with end user 12 may register their presence at a remote location through

15 foreign agents 54a-n. Any one of foreign agents 54a-n may communicate with selected home agents 60a-n such that data packets may be appropriately forwarded to their proper destination. The use of foreign agents 54a-n allows end user 12 to freely roam from one network

20 coverage area to another while maintaining a communication session and retaining its identity (as illustrated in the example provided below). Foreign agents 54a-n may be positioned in any suitable location within or external to communication system 10 according

25 to particular needs. Foreign agents 54a-n and home agents 60a-n may represent tunnel initiators or tunnel terminators being provided to communication system 10. In other embodiments, these elements may be substituted with any element that operates to initiate or to provide

30 a contact for communication sessions in a network environment.

In operation of an example embodiment, end user 12 may establish a connection to a selected one of foreign agents 54a-n. End user 12 may have roamed into an area covered by foreign agent 54a or 54b (as illustrated by a hatched line 42 in FIGURE 3 showing movement of end user 12 while maintaining a connectivity to access network 22). When end user 12 roams from one foreign agent to another, the associated IP address of end user 12 (or his/her identity) does not require a change from the perspective of other elements in the network. This may be generally accomplished by storing the unique identity or IP address with a selected home agent 60a-n.

In the case of mobile IP communications, a mobile IP tunnel may then be established between foreign agent 54a and loadbalancer 26 using a request message during the tunnel establishment phase. This is illustrated in FIGURE 3 as a communications pathway (or link) 34a. Loadbalancer 26 may then forward the message to a selected home agent 60a-n. The selected home agent 60a-n may respond to receiving the request by positioning an identification element (such as its associated IP address) in the response that gets communicated back through loadbalancer 26. Alternatively, any other element in communication system 10 may position the identification element in the response.

After the establishment phase is complete, traffic may be sent directly to a selected home agent 60a-n from foreign agent 54a or foreign agent 54b in cases where end user 12 has roamed. The communication between foreign agent 54b and loadbalancer 26 is illustrated by a communications pathway (or link) 34b. With the initiation step complete, destination server 38 may send

communications or data packets through the selected home agent 60a-n and to a selected foreign agent 54a or 54b. This pattern illustrates asymmetrical routing, whereby loadbalancer 26 is implicated in the establishment of a communication session or tunnel and then removed from the communication flow during data transfer. This feature minimizes the involvement of loadbalancer 26 and provides for more direct communications between foreign agents 54a and 54b and selected home agents 60a-n, as illustrated by a set of dashed lines 28a and 28b in FIGURE 3.

In cases where end user 12 roams from foreign agent 54a to foreign agent 54b, foreign agent 54b may establish a mobile IP tunnel (in the case of mobile IP communications) to the selected home agent 60a-n. The selected home agent 60a-n may then respond by communicating requested data from destination server 38 to foreign agent 54b in a seamless or transparent fashion. This seamless handover or handoff may be executed such that the communication session is uninterrupted and properly maintained during all data exchanges involving end user 12.

FIGURE 4 is a flowchart illustrating a series of steps associated with a method for communicating in a loadbalancing environment. The method begins at step 100 where a request packet is generated by end user 12 in order to initiate a communication session or tunnel. At step 102, the request packet is communicated through loadbalancer 26 and to a node in the network (such as a selected home agent 60a-n or a selected TPNS 30a-n) in order to establish a communication session or tunnel. At step 104, end user 12 may be properly authenticated

and/or authorized by invoking an element such as AAA server 32.

At step 106, a response is generated and communicated through loadbalancer 26 that includes an
5 identification element (such as an IP address or destination information associated with the node in the network that established the communication session, e.g. the TPNS or the home agent). At step 108, data transfer may be executed between end user 12 and the node in the
10 network that established the communication session such that loadbalancer 26 may not be implemented in a communication session or stream associated with the data transfer.

Some of the steps illustrated in FIGURE 4 may be
15 changed or deleted where appropriate and additional steps may also be added to the flowchart. These changes may be based on specific communications architectures or particular interfacing arrangements and configurations of associated elements and do not depart from the scope or
20 the teachings of the present invention.

Although the present invention has been described in detail with reference to mobile IP communications (in conjunction with the use of home agents and foreign agents) and NAS 18 and TPNSs 30a-n, communication system
25 10 may be used for any tunneling protocol involving a redirection or handoff of communications in a network environment. Any suitable communications that involve the transitioning between an initialization state and a data transfer state may benefit from the teachings of the
30 present invention. Mobile IP applications and TPNS communications have only been offered for purposes of

teachings and should not be construed to limit the scope of the present invention in any way.

In addition, communication system 10 may be extended to any scenario in which end user 12 is provided with mobility (in the context of a wired or a wireless connection or coupling) and communicates with some type of access server (e.g. NAS 18, foreign agents 54a-n, etc.). End user 12 may be using a dedicated connection of some form, or using forms of multiple access protocols where appropriate. Access may be associated with PPP or alternatively with layer three protocol (for example IP) over an L2 layer in accordance with particular needs. Such an embodiment would include any suitable tunnel terminators and/or tunnel initiators that may be operable to communicate with loadbalancer 26.

Numerous other changes, substitutions, variations, alterations, and modifications may be ascertained by those skilled in the art and it is intended that the present invention encompass all such changes, substitutions, variations, alterations, and modifications as falling within the spirit and scope of the appended claims.

Moreover, the present invention is not intended to be limited in any way by any statement in the specification that is not otherwise reflected in the appended claims. Various example embodiments have been shown and described, but the present invention is not limited to the embodiments offered. Accordingly, the scope of the present invention is intended to be limited solely by the scope of the claims that follow.

WHAT IS CLAIMED IS:

1. A method for communicating in a loadbalancing environment, the method comprising:

receiving a request packet from a network access
5 server (NAS) to initiate a communication session;

communicating the request packet to a tunneling protocol network server (TPNS); and

receiving a response packet in response to the request packet, the response packet establishing a tunnel
10 that facilitates the communication session and that includes an identification element associated with the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS.

15

2. The method of Claim 1, further comprising:

authenticating an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20

3.. The method of Claim 1, wherein the identification element includes an internet protocol (IP) address associated with the TPNS such that the NAS may communicate the data transfer to the TPNS.

25

4. The method of Claim 1, wherein the data transfer between the TPNS and the NAS originates from a server that is coupled to a network and that is sought to be accessed by the end user.

5. The method of Claim 1, further comprising:
providing a communications platform for the
communication session and the tunnel with an access
network, the access network operable to provide a
5 communicative interface between the NAS and the TPNS.

6. The method of Claim 1, wherein the request
packet is communicated through a radio access network
(RAN) and the NAS before reaching a loadbalancer that
10 facilitates establishment of the tunnel.

7. A method for communicating in a loadbalancing environment, the method comprising:

receiving a request packet from a foreign agent to initiate a communication session;

5 communicating the request packet to a home agent; and

receiving a response packet in response to the request packet, the response packet establishing a tunnel that facilitates the communication session and that
10 includes an identification element associated with the home agent such that a data transfer associated with the communication session is executed between the foreign agent and the home agent.

15 8. The method of Claim 7, further comprising:

authenticating an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20 9. The method of Claim 7, wherein the identification element includes an internet protocol (IP) address associated with the home agent such that the foreign agent may communicate the data transfer to the home agent.

10. The method of Claim 7, further comprising:

identifying movement by an end user associated with the communication session from a first service area to a second service area; and

5 directing the data transfer to an additional foreign agent associated with the second service area such that the communication session is handed off between foreign agents while maintaining the communication session.

10 11. The method of Claim 7, wherein the data transfer between the home agent and the foreign agent originates from a server that is coupled to a network and that is sought to be accessed by an end user associated with the communication session.

15

12. The method of Claim 7, further comprising:

providing a communications platform for the communication session and the tunnel with an access network, the access network operable to provide a
20 communicative interface between the foreign agent and an end user associated with the communication session.

13. A system for communicating in a loadbalancing environment, comprising:

means for receiving a request packet from a network access server (NAS) to initiate a communication session;

5 means for communicating the request packet to a tunneling protocol network server (TPNS); and

means for receiving a response packet in response to the request packet, the response packet establishing a tunnel that facilitates the communication session and
10 that includes an identification element associated with the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS.

15 14. The system of Claim 13, further comprising:

means for authenticating an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20 15. The system of Claim 13, wherein the identification element includes an internet protocol (IP) address associated with the TPNS such that the NAS may communicate the data transfer to the TPNS.

25 16. The system of Claim 13, wherein the data transfer between the TPNS and the NAS originates from a server that is coupled to a network and that is sought to be accessed by the end user.

17. The system of Claim 13, further comprising:
means for providing a communications platform for
the communication session and the tunnel with an access
network, the access network operable to provide a
5 communicative interface between the NAS and the TPNS.

18. A system for communicating in a loadbalancing environment, comprising:

means for receiving a request packet from a foreign agent to initiate a communication session;

5 means for communicating the request packet to a home agent; and

means for receiving a response packet in response to the request packet, the response packet establishing a tunnel that facilitates the communication session and
10 that includes an identification element associated with the home agent such that a data transfer associated with the communication session is executed between the foreign agent and the home agent.

15 19. The system of Claim 18, further comprising:

means for authenticating an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20 20. The system of Claim 18, wherein the identification element includes an internet protocol (IP) address associated with the home agent such that the foreign agent may communicate the data transfer to the home agent.

21. The system of Claim 18, further comprising:

means for identifying movement by an end user associated with the communication session from a first service area to a second service area; and

5 means for directing the data transfer to an additional foreign agent associated with the second service area such that the communication session is handed off between foreign agents while maintaining the communication session.

10

22. The system of Claim 18, wherein the data transfer between the home agent and the foreign agent originates from a server that is coupled to a network and that is sought to be accessed by an end user associated
15 with the communication session.

23. The system of Claim 18, further comprising:

means for providing a communications platform for the communication session and the tunnel with an access
20 network, the access network operable to provide a communicative interface between the foreign agent and an end user associated with the communication session.

24. Software for communicating in a loadbalancing environment that is embodied in a computer readable media and operable to:

5 receive a request packet from a network access server (NAS) to initiate a communication session;

 communicate the request packet to a tunneling protocol network server (TPNS); and

 receive a response packet in response to the request
10 packet, the response packet establishing a tunnel that facilitates the communication session and that includes an identification element associated with the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS.

15

25. The software of Claim 24, further operable to:

 authenticate an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20

26. The software of Claim 24, wherein the identification element includes an internet protocol (IP) address associated with the TPNS such that the NAS may communicate the data transfer to the TPNS.

25

27. The software of Claim 24, further operable to:

 provide a communications platform for the communication session and the tunnel with an access network, the access network operable to provide a
30 communicative interface between the NAS and the TPNS.

28. Software for communicating in a loadbalancing environment that is embodied in a computer readable media and operable to:

5 receive a request packet from a foreign agent to initiate a communication session;

communicate the request packet to a home agent; and

10 receive a response packet in response to the request packet, the response packet establishing a tunnel that facilitates the communication session and that includes an identification element associated with the home agent such that a data transfer associated with the communication session is executed between the foreign agent and the home agent.

15 29. The software of Claim 28, further operable to:

authenticate an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20 30. The software of Claim 28, wherein the identification element includes an internet protocol (IP) address associated with the home agent such that the foreign agent may communicate the data transfer to the home agent.

25 31. The software of Claim 28, further operable to:

identify movement by an end user associated with the communication session from a first service area to a second service area; and

30 direct the data transfer to an additional foreign agent associated with the second service area such that the communication session is handed off between foreign agents while maintaining the communication session.

32. An apparatus for communicating in a loadbalancing environment, comprising:

a loadbalancer operable to receive a request packet from a network access server (NAS) operable to generate the request packet that initiates a communication session, wherein a tunneling protocol network server (TPNS) is operable to receive the request packet from the loadbalancer and in response to the request packet generate a response packet, the response packet establishing a tunnel that facilitates the communication session and that includes an identification element associated with the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS.

15

33. The apparatus of Claim 32, further comprising:

an authentication, authorization, and accounting (AAA) server operable to authenticate an end user associated with the communication session before establishing the tunnel that facilitates the communication session.

20

34. The apparatus of Claim 32, wherein the identification element includes an internet protocol (IP) address associated with the TPNS such that the NAS may communicate the data transfer to the TPNS.

25

35. The apparatus of Claim 32, wherein the data transfer between the TPNS and the NAS originates from a server that is coupled to a network and that is sought to be accessed by the end user.

30

36. The apparatus of Claim 32, further comprising:
an access network operable to provide a
communications platform for the communication session and
the tunnel, wherein the access network is further
5 operable to provide a communicative interface between the
NAS and the TPNS.

37. An apparatus for communicating in a loadbalancing environment, comprising:

a loadbalancer operable to receive a request packet from a foreign agent, the foreign agent being operable to
5 generate the request packet in order to initiate a communication session, wherein a home agent is operable to receive the request packet and in response to the request packet generate a response packet, the response packet establishing a tunnel that facilitates the
10 communication session and that includes an identification element associated with the home agent such that a data transfer associated with the communication session is executed between the foreign agent and the home agent.

15 38. The apparatus of Claim 37, further comprising:
an authentication, authorization, and accounting (AAA) server operable to authenticate an end user associated with the communication session before
establishing the tunnel that facilitates the
20 communication session.

39. The apparatus of Claim 37, wherein the identification element includes an internet protocol (IP) address associated with the home agent such that the
25 foreign agent may communicate the data transfer to the home agent.

40. The apparatus of Claim 37, further comprising:
an access network operable to provide a
communications platform for the communication session and
the tunnel, wherein the access network is further
5 operable to provide a communicative interface between the
foreign agent and an end user associated with the
communication session.

41. An apparatus for communicating in a loadbalancing environment, comprising:

a network access server (NAS) operable to generate a request packet that initiates a communication session;

5 a loadbalancer operable to receive the request packet from the NAS;

a tunneling protocol network server (TPNS) operable to receive the request packet from the loadbalancer and in response to the request packet generate a response
10 packet, the response packet establishing a tunnel that facilitates the communication session and that includes an identification element associated with the TPNS such that a data transfer associated with the communication session is executed between the NAS and the TPNS, wherein
15 the identification element includes an internet protocol (IP) address associated with the TPNS such that the NAS may communicate the data transfer to the TPNS; and

an authentication, authorization, and accounting (AAA) server operable to authenticate an end user
20 associated with the communication session before establishing the tunnel that facilitates the communication session.

1/2

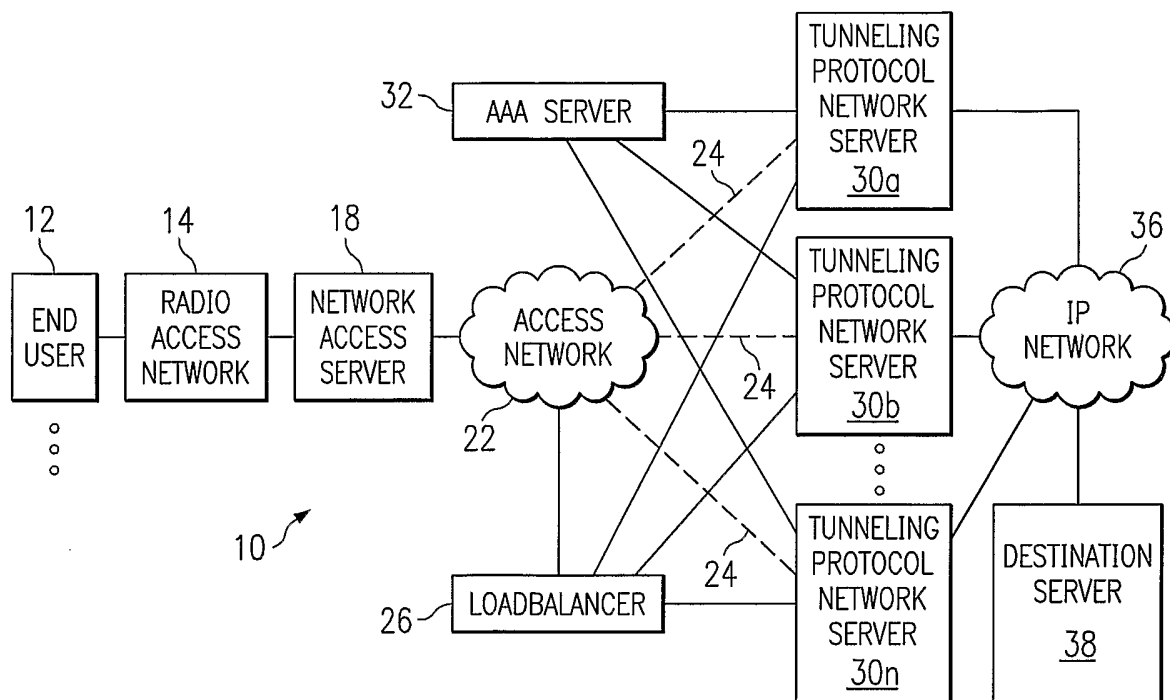


FIG. 1

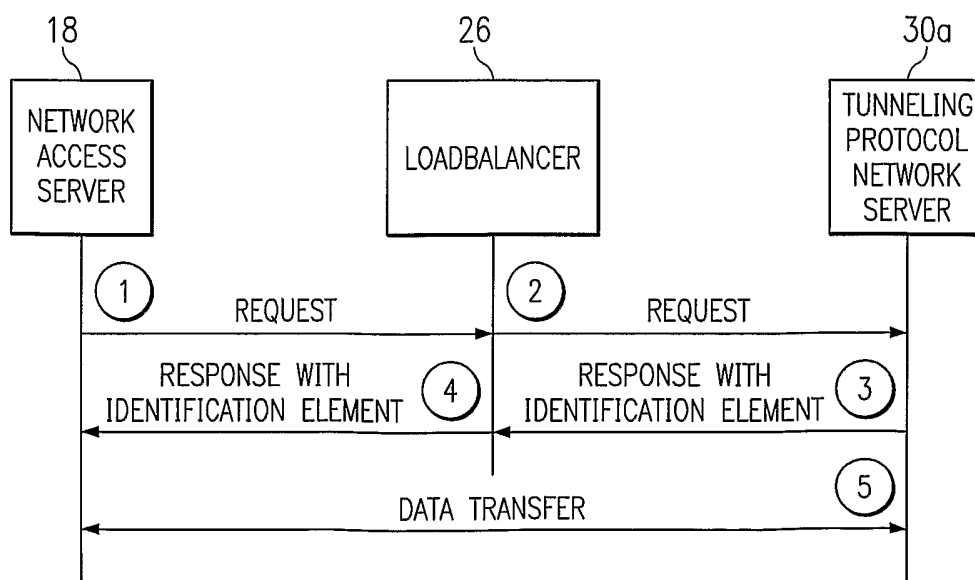


FIG. 2

