

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2008311812 B2**

(54) Title
Method and system for real-time media synchronisation across a network

(51) International Patent Classification(s)
H04L 29/06 (2006.01) **H04L 12/58** (2006.01)
H04L 1/00 (2006.01)

(21) Application No: **2008311812** (22) Date of Filing: **2008.10.17**

(87) WIPO No: **WO09/052428**

(30) Priority Data

(31) Number	(32) Date	(33) Country
12/192,890	2008.08.15	US
60/999,619	2007.10.19	US
12/028,400	2008.02.08	US
61/093,278	2008.08.29	US

(43) Publication Date: **2009.04.23**

(44) Accepted Journal Date: **2013.01.17**

(71) Applicant(s)
Voxer IP LLC

(72) Inventor(s)
Ranney, Matthew J.;Panttaja, James T.;Panttaja, Mary G.;Katis, Thomas E.

(74) Agent / Attorney
Shelston IP, L 21 60 Margaret St, Sydney, NSW, 2000

(56) Related Art
US 7315898 B2
EP 1725060 A1

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
23 April 2009 (23.04.2009)

PCT

(10) International Publication Number
WO 2009/052428 A1

(51) International Patent Classification:

H04L 29/06 (2006.01) **H04L 1/00** (2006.01)
H04L 12/58 (2006.01)

(21) International Application Number:

PCT/US2008/080369

(22) International Filing Date: 17 October 2008 (17.10.2008)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/999,619	19 October 2007 (19.10.2007)	US
12/028,400	8 February 2008 (08.02.2008)	US
12/192,890	15 August 2008 (15.08.2008)	US
61/093,278	29 August 2008 (29.08.2008)	US

(71) Applicant (for all designated States except US): **RE-BELVOX, LLC** [US/US]; 512 Second Street, San Francisco, California 94107 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KATIS, Thomas, E.**

[US/US]; 970 East Broadway, E-321, Jackson, Wyoming 83001-9475 (US). **PANTTAJA, James, T.** [US/US]; 4016 West Soda Rock Lane, Healdsburg, California 95448 (US). **PANTTAJA, Mary, G.** [US/US]; 4016 West Soda Rock Lane, Healdsburg, California 95448 (US). **RANNEY, Matthew, J.** [US/US]; 280 Euclid Avenue, Oakland, California 94610 (US).

(74) Agents: **BEYER, Steve, D.** et al.; BEYER LAW GROUP LLP, P. O. Box 1687, Cupertino, California 95015-1687 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR REAL-TIME MEDIA SYNCHRONISATION ACROSS A NETWORK

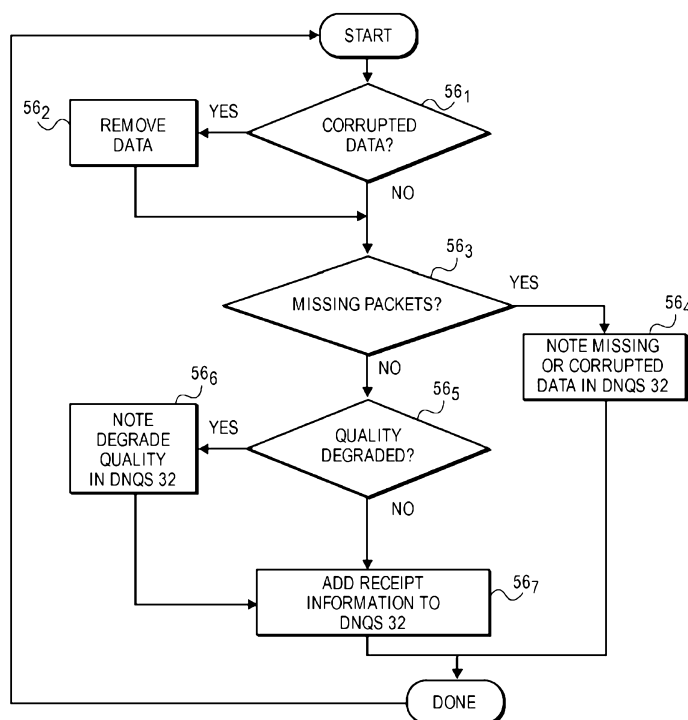


FIG. 9D

(57) Abstract: Methods and systems for progressive synchronization in real-time of stored copies of media transmitted between nodes on a network. The methods and systems include progressively transmitting available indexed media from a sending node to a receiving node with a packet size and packetization interval sufficient to enable the near real-time rendering of the indexed media, wherein the near real-time rendering of the indexed media provides a recipient with an experience of reviewing the transmitted media live. At the receiving node, the transmitted indexed media is progressively receive and any indexed-media that is not already locally stored at the receiving node is noted. The methods and systems also include the real-time synchronization of the media of a conversation between two communication devices over the network by requesting missing media from the sending node. As voice or other media is created at either communication device, the copies of the media are synchronized in real-time across the network. The methods and systems also include optimized routing protocols that are simple to implement and decrease network traffic and latency.



(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

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

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

METHOD AND SYSTEM FOR REAL-TIME MEDIA SYNCHRONISATION ACROSS A NETWORK

BACKGROUND

Field of the Invention.

[0001] This invention pertains to telecommunications. In more particular embodiments, this invention pertains to a method and apparatus for the near real-time synchronization of voice communications.

Description of Related Art

[0002] Any discussion of the prior art throughout the specification should in no way be considered as an admission that such prior art is widely known or forms part of common general knowledge in the field.

[0002a] The current state of voice communications suffers from inertia. In spite of automated switching, high bandwidth networks and technologies such as satellites, fiber optics, Voice over IP (VoIP), wireless and cellular networks, there has been little change in how people use telephones. One is still required to pick up the phone, dial another party, wait for a connection to be made, and then engage in a full-duplex, synchronous conversation with the dialed party. If the recipient does not answer, no connection is made, and the conversation does not take place.

[0003] At best, a one-way asynchronous voice message may be left if the recipient has voice mail. The process of delivering the voice mail, however, is burdensome and time consuming. The caller is required to wait for the phone on the other end to stop ringing, transition into the voice mail system, listen to a voice message greeting, and then leave the message. Current voice mail systems are also inconvenient for the recipient. The recipient has to dial a code to access their voice mail, navigate through a series of prompts, listen to any earlier received voice messages in the queue, and then finally listen to the message of the sender.

[0004] Another drawback with typical voice mail systems is the inability to organize or permanently archive voice messages. With some voice mail systems, a user may save a message, but it is automatically deleted after a predetermined period of time and lost forever.

[0005] Yet another problem with current voice mail systems is that a connection must be made between the caller and the voice mail system before a message can be left. If no connection is made, there is no way for the caller to leave a message.

[0006] Current telephone systems are based on relatively simplistic usage patterns: real-time live calls or disjointed voice mail messages, which are typically deleted as

they are heard. These forms of voice communications do not capture the real power that can be achieved with voice communication or take advantage of the advances of network speed and bandwidth that is now available. Also, if the phone network is down, or is inaccessible, (e.g., a cell phone user is in an area of no coverage or the
5 phone lines are down due to bad weather), no communication can take place.

[0007] In general, telephone based communications have not kept pace with the advances in text-based communications. Instant messaging, emailing, faxing, chat groups, and the ability to archive text messages, are all commonplace with text based communications. Other than voice mail, there are few existing tools available to
10 manage and/or archive voice messages. In comparison, the tools currently available to manage telephone communications are primitive compared to text communications.

[0008] The corporate environment provides just one example of the weakness in current voice communication tools. There is currently no integrated way to manage voice communications across an organization as a corporate asset. Employees
15 generally do not record or persistently store their phone conversations. Most business related voice communication assets are gone as quickly as the words are spoken, with no way to manage or store the content of those conversations in any manageable form.

[0009] As an illustrative example, consider a sales executive at a company. During
20 the course of a busy day, the executive may make a number of calls, and close several sales, with customers over the phone. Without the ability to organize, store, and later retrieve these conversations, there is no way for the executive to resolve potential issues that may arise, such as recalling the terms of one deal versus another, or challenging a customer who disputes the terms of a previously agreed upon sale. If
25 this executive had the ability to easily retrieve and review conversations, these types of issues could be easily and favorably resolved.

[0010] Current tactical radio systems, such as those used by the military, fire, police, paramedics, rescue teams, and first responders, also suffer from a number of deficiencies. Most tactical radio communication must occur through a "live" radio
30 connection between the sender of a message and a recipient. If there is no radio connection between the two parties, there can be no communication. Urgent messages cannot be sent if either the sender or the receiver does not have access to their radio, or a radio circuit connection is not established. Tactical communications are therefore

plagued with several basic problems. There is no way (i) to guarantee the delivery of messages, (ii) for a recipient to go back and listen to a message that was not heard in real time; (iii) to control the granularity of the participants in a conversation; (iv) for the system to cope when there is a lack of signal integrity for a live conversation. If a message is not heard live, it is missed. There are no tools for either the sender or a recipient to manage, prioritize, archive and later retrieve (i.e. time-shift) the messages of a conversation that were previously sent.

[0011] Yet another drawback with tactical radio communication systems is that only one message can be sent at a time per channel. Consider an example of a large building fire, where multiple teams of fire fighters, police and paramedics are simultaneously rescuing victims trapped in the building, fighting the fire, providing medical aid to victims, and controlling bystanders. If each of the teams is using the same channel, communications may become crowded and chaotic. Transmissions get “stepped on” when more than one person is transmitting at the same time. Also there is no way to differentiate between high and low priority messages. A team inside the burning building fighting the fire or rescuing trapped victims should have a higher priority over other teams, such as those controlling bystanders. If high priority messages are stepped on by lower priority messages, it could not only hamper important communications, but could endanger the lives of the fire fighters and victims in the building.

[0012] One possible solution to the lack of ability to prioritize messages is to use multiple channels, where each team is assigned a different channel. This solution, however, creates its own set of problems. How does the fire chief determine which channel to listen too at any point in time? How do multiple teams communicate with one another if they are all on different channels? If one team calls for urgent help, how are other teams to know if they are listening to other channels? While multiple channels can alleviate some issues, it can also cause confusion, creating more problems than if a single channel is used.

[0013] The lack of management tools that effectively prioritize messages, that allow multiple conversations to take place at the same time, that enable the time-shifting of messages to guarantee delivery, or that support archiving and storing conversations for later retrieval and review, all contribute to the problems associated with tactical radios. In first responder situations, such as with the military, police, and fire,

effective communication tools can literally mean the difference between life and death, or the success or failure of a mission. The above burning building example is useful in illustrating just some of the issues with current tactical radio communications. Similar problems exist with the military, police, first responders and
5 others who use tactical communications.

[0014] With packet-based networks, commonly used protocols include Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). UDP offers the advantage of fast delivery of data, but at the expense of completeness. Packets may be dropped in transit and not available when attempting to render the data as soon as
10 possible at the destination. In spite of the shortcomings, UDP is the standard for Voice over Internet Protocol (VoIP) transmissions due to its speed attributes. On the other hand TCP does guarantee the delivery of perfect (i.e., an exact copy of the transmitted data) data, but at the expense of latency. All packets are delivered, regardless of how long it takes. This delay makes TCP impractical for use with “live”
15 phone calls. Currently there are no known protocols that offer the performance advantages of both TCP and UDP, where “good enough” media can be transmitted for rendering as soon as possible, with the eventual delivery of a perfect copy of the media. Also there is no protocol that determines how much information should be sent over the network based on the presence of recipients on the network and their
20 intentions to render the data either live or in a time-shifted mode. In addition, other factors commonly considered, such as network latency, network degradation, packet loss, packet damage, and general bandwidth conditions, are used in determining how much data to transmit. Prior art systems, however, do not consider the presence and intentions of recipients. As a result, the default assumption is that the data is rendered
25 by the recipient in real time. When a recipient is not going to render data immediately, these prior art systems unnecessarily use bandwidth when it is not needed, degrading the overall performance of the network.

[0015] For the reasons recited above, telephone, voicemail and tactical voice communications systems are inadequate. An improved voice and media
30 communication and management system and method, and improvements in delivering voice and other media over packet-based networks, including the near real-time synchronization of the storage of voice and other media at both the sending and

receiving communication devices and along each hop on the network between the two communication devices is therefore needed.

SUMMARY OF THE INVENTION

[0016] Embodiments of the present invention are directed to methods and systems for progressive synchronization in real-time of stored copies of media transmitted between nodes on a network. The methods and systems include progressively transmitting available indexed media from a sending node to a receiving node with a packet size and packetization interval sufficient to enable the near real-time rendering of the indexed media, wherein the near real-time rendering of the indexed media provides a recipient with an experience of reviewing the transmitted media live. At the receiving node, the transmitted indexed media is progressively receive and any indexed media that is not already locally stored at the receiving node is noted. The methods and systems also include the real-time synchronization of the media of a conversation between two communication devices over the network. As voice or other media is created at either communication device, the copies of the media are synchronized in real-time across the network. The methods and systems also include optimized routing protocols that are simple to implement and decrease network traffic and latency.

[0016a] According to a first aspect of the present invention there is provided a method for progressively synchronizing stored copies of voice media transmitted over a network during a conversation that may be selectively conducted in a real-time mode and in a time-shifted mode, comprising:

- (a) progressively receiving voice media transmitted from a sending node to a receiving node as the voice media is persistently stored at the sending node, the voice media transmitted with a packet size and packetization interval sufficient to enable the near real-time rendering of the voice media;

- (b) progressively and persistently storing the transmitted voice media at the receiving node as the voice media is received;

- (c) noting any voice media that was transmitted but not received and persistently stored at the receiving node;

- (d) generating requests as needed for any noted voice media;

- (e) receiving retransmissions of the noted voice media from the sending node to the receiving node in response to the requests;

- (f) repeating (d) and (e) until the sending node and the receiving node each have complete copies of the voice media maintained in persistent storage; and

(g) selectively rendering the transmitted voice media at the receiving node either in the real-time mode as the voice media is received over the network or out of persistent storage in the time-shifted mode.

[0016b] According to a second aspect of the present invention there is provided a programmable communication device configured to carry out the method of the first aspect.

[0016c] According to a third aspect of the present invention there is provided a computer readable medium bearing an application executable by a data processor of a communication device for configuring the communication device to carry out the method of the first aspect.

[0016d] According to a fourth aspect of the present invention there is provided a method performed on a client communication device, comprising:

synchronizing copies of persistently stored transmitted and received voice messages exchanged between the client communication device and a remote client communication device, the synchronization performed until both the client communication device and the remote client communication device have persistently stored complete copies of both the voice messages transmitted by the client communication device to the remote client communication device and the voice messages received by the client communication device from the remote client communication device, wherein the synchronization occurs and completes, if necessary, beyond a time when it is not possible to render at least one of the transmitted and received voice messages in real-time; and

selectively rendering the voice media of the received messages on the client communication device in both:

(i) a real-time mode as the voice media is received from the remote client communication device; and

(ii) a time-shifted mode by retrieving and rendering the received voice media from persistent storage,

wherein the synchronization occurs and completes, as necessary, beyond a time at which it is no longer possible to render the voice media in real-time, and

wherein persistently stored means for a time period beyond what is needed for the client communication device to:

transmit the voice media of the transmitted messages; and

render the received voice media of the received voice messages in the real-time mode.

[0016e] According to a fifth aspect of the present invention there is provided a programmable communication device configured to carry out the method of the fourth aspect.

[0016f] According to a sixth aspect of the present invention there is provided a computer readable medium bearing an application executable by a data processor of a communication device for configuring the communication device to carry out the method of the fourth aspect.

[0016g] It is an object of the present invention to overcome or ameliorate at least one of the disadvantages of the prior art, or to provide a useful alternative.

[0016h] Unless the context clearly requires otherwise, throughout the description and the claims, the words “comprise”, “comprising”, and the like are to be construed in an inclusive sense as opposed to an exclusive or exhaustive sense; that is to say, in the sense of “including, but not limited to”.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] The invention may best be understood by reference to the following description taken in conjunction with the accompanying drawings, which illustrate specific embodiments of the invention.

[0018] Figure 1 is a diagram of the architecture of the communication and media management system of the invention.

[0019] Figures 2A and 2B illustrate a block diagram of a Client running on a Device in the communication and management system of the invention.

[0020] Figure 3 is a block diagram of a Server used in the communication and media management system of the invention.

[0021] Figures 4A through 4D illustrate various embodiments of data payloads used in the communication and management system of the invention.

[0022] Figure 5 is a diagram illustrating data being transmitted over a shared IP network in accordance with the invention.

[0023] Figure 6 is a diagram illustrating data being transmitted over a circuit-based network in accordance with the invention.

[0024] Figure 7 is a diagram illustrating data being transmitted across both a cellular network and the Internet in accordance with the invention.

[0025] Figures 8A through 8F are a series of flow diagrams illustrating a store and stream function of the communication and management system of the invention.

5 [0026] Figures 9A through 9C are flow diagrams illustrating the operation of a Payload Quality Manager (PQM) and Figures 9D through 9F are flow diagrams illustrating the Data Quality manager (DQM), both used by the Clients and Servers of the invention.

10 [0027] Figure 10 is a diagram illustrating the near real-time synchronization of the media of a conversation between two communication devices in accordance with the present invention.

[0028] Figure 11 is a flow diagram illustrating the sequence for the real-time synchronization of indexed media between two communication devices according to the present invention.

15 [0029] Figure 12 is an example of a distributed services architecture according to the present invention.

[0030] Figure 13 is another example of a distributed services architecture for illustrating a routing protocol according to the present invention.

20 [0031] Figure 14 is another example of another distributed services architecture for showing near real-time synchronization of media at each hop across the network according to the present invention.

[0032] Figures 15A and 15B are block diagrams that illustrate the hardware used for running the Client and Server applications of the invention.

25 [0033] It should be noted that like reference numbers refer to like elements in the figures.

DETAILED DESCRIPTION OF SPECIFIC EMBODIMENTS

[0034] The invention will now be described in detail with reference to various embodiments thereof as illustrated in the accompanying drawings. In the following description, specific details are set forth in order to provide a thorough understanding
30 of the invention. It will be apparent, however, to one skilled in the art, that the invention may be practiced without using some of the implementation details set forth herein. It should also be understood that well known operations have not been described in detail in order to not unnecessarily obscure the invention.

A. Functional Overview

[0035] The communication media management method and system supports new modes of engaging in voice conversations and/or managing multiple simultaneous conversations using a variety of media types, such as voice, video, text, location, sensor information, and other data. Users can engage in conversations by sending voice messages to designated recipients. Depending on preferences and priorities, the recipient(s) might participate in the conversation in real time, or they might simply be notified that the message is ready for retrieval. In the latter case, the recipient participates in the conversation in a time-shifted mode by reviewing and replying to the recorded message at their convenience.

[0036] Users are empowered to conduct communications in either: (i) a near-synchronous or “live” conversation, providing a user experience similar to a standard full duplex phone call; or (ii) in a series of back and forth time-delayed transmissions (i.e., time-shifted mode). Further, users engaged in a conversation can seamlessly transition from the live mode to the time-shifted mode and back again. This attribute also makes it possible for users to engage in multiple conversations, at the same time, by prioritizing and shifting between the two modes for each conversation. Two individuals using the system can therefore send recorded voice messages back and forth to each other and review the messages when convenient, or the messages can be sent at a rate where they essentially merge into a live, synchronous voice conversation. This new form of communication, for the purposes of the present application, is referred to as “Voxing”

[0037] When you “Vox” someone, the conversation consists of a series of discrete recorded messages, which are recorded in a number of locations, which may include the encoding device of the sender, (e.g. a phone or computer), servers on multiple transmission hops across the network, and the receiver's rendering device. Unlike a standard phone call or voice mail, the system provides the following features and advantages: (i) the conversation can transition between live and time-shifted or vice versa; (ii) the discrete messages of the conversation are semantically threaded together and archived; (iii) since the messages are recorded and are available for later retrieval, attention can be temporarily diverted from the conversation and then the conversation can be later reviewed when convenient; (iv) the conversation can be paused for seconds, minutes, hours, or even days, and can be picked up again where

left off; (v) one can rejoin a conversation in progress and rapidly review missed messages and catch up to the current message (i.e., the live message); (vi) no dedicated circuit is needed for the conversation to take place, as required with conventional phone calls; and (vii) lastly, to initiate a conversation, one can simply
5 begin transmitting to an individual or a group. If the person or persons on the other end notice that they are receiving a message, they have the option of reviewing and conducting a conversation in real time, or reviewing at a later time of their choice.

[0038] The communication media management system also supports new modes of optimizing the transmission of data over a network. The system actively manages the
10 delivery of payloads to a recipient engaged in a conversation in real time when network conditions are less than ideal. For example when network conditions are poor, the system intentionally reduces the quality of the data for transmission to the point where it is “good enough” to be rendered upon receipt by the recipient, allowing the real time participation of the conversation. The system also guarantees the
15 eventual delivery of an “exact” copy of the messages over time. The system and method therefore provides the advantages of both speed and accuracy. The utilization of network bandwidth is optimized by making tradeoffs between timeliness and media quality, using the presence and intentions of whether or not recipient(s) intend to review to the message immediately in real time, as well as measures of network
20 latency, network degradation, packet loss or damage, and/or current bandwidth conditions.

[0039] It should be noted that the messages of conversations may contain voice only or voice, video and other data, such as sensor information. When the messages are reviewed, they are listened to or visually reviewed, or a combination thereof,
25 depending on the type of media contained in the messages. Although as of the filing of the present application, most conversations are voice only, it is intended that the communication system and method described herein broadly includes conversations including multiple media types, such as voice and video for example.

[0040] An improved voice and other media communication and management system
30 and method is disclosed which provides one or more of the following features and functions:

- i. enabling users to participate in multiple conversation types, including live phone calls, conference calls, voice messaging, consecutive or simultaneous communications;
- ii. enabling users to review the messages of conversations in either a live mode or a
5 time-shifted mode (voice messaging);
- iii. enabling users to seamlessly transition a conversation between a synchronous “live” mode and a time shifted mode;
- iv. enabling users to participate in conversations without waiting for a connection to
10 be established with another participant or the network. This attribute allows users to begin conversations, participate in conversations, and review previously received time-shifted messages of conversations even when there is no network available, when the network is of poor quality, or other participants are unavailable;
- v. enabling the system to save media payload data at the sender and, after network
15 transmission, saving the media payload data at all receivers;
- vi. enabling the system to organize messages by threading them sequentially into semantically meaningful conversations in which each message can be identified and tied to a given participant in a given conversation;
- vii. enabling users to manage each conversation with a set of user controlled
20 functions, such as reviewing “live”, pausing or time shifting the conversation until it is convenient to review, replaying in a variety of modes (e.g., playing faster, catching up to live, jump to the head of the conversation) and methods for managing conversations (archiving, tagging, searching, and retrieving from archives);
- 25 viii. enabling the system to manage and share presence data with all conversation participants, including online status, intentions with respect to reviewing any given message in either the live or time-shifted mode, current attention to messages, rendering methods, and network conditions between the sender and receiver;
- 30 ix. enabling users to manage multiple conversations at the same time, where either (a) one conversation is current and all others are paused; (b) multiple conversations are rendered consecutively, such as but not limited to tactical communications; or

- (c) multiple conversations are active and simultaneously rendered, such as in a stock exchange or trading floor environment.
- x. enabling users to store all conversations, and if desired, persistently archive them in a tangible medium, providing an asset that can be organized indexed, searched, transcribed, translated and/or reviewed as needed;
- xi. enabling the system to provide real time call functionality using a best-efforts mode of message delivery at a rate “good enough” for rendering as soon as possible (similar to UDP), and the guaranteed eventual delivery of exact copies of the messages as transmitted by requesting retransmission of any missing or defective data from the originally saved perfect copy (similar to TCP); and
- xii. enabling the system to optimize the utilization of network bandwidth by making tradeoffs between timeliness and media quality, using the presence and intentions of the recipient(s) (i.e., to either review the media in real-time or in a time-shifted mode), as well as measures of network latency, network degradation, packet loss or damage, and/or current bandwidth conditions.

[0041] In various embodiments, some or all of the numerous features and functions listed above may be implemented. It should be understood, however, that different embodiments of the invention need not incorporate all of the above listed features and functions.

B. Glossary

[0042] Prior to explaining the details of the invention, it is useful to define some of the terms and acronyms used throughout the written description. This glossary of terms is organized into groups of System Components, Media, Media Management, People and Conversation Management.

B.1. System Components

[0043] **Client:** A Client is the user application in the communication system, which includes a user interface, persistent data storage, and “Voxing” functionality. Users interact with the Client application, and the Client application manages all communications (messages and signals) and payload (Media) transfers that are transmitted or received over a network. The Client supports encoding of media (e.g., the capturing of voice, video, or other data content) and the rendering of media and supports security, encryption and authentication as well as the optimization of the

transmission of data across the network. A Client may be used by one or multiple Users (i.e., multi-tenant).

[0044] Device: A physical device that runs the Client application. A User may be actively logged into a single Device or multiple Devices at any given point of time.

5 In various embodiments, a Device may be a general-purpose computer, a portable computing device, a programmable phone, a programmable radio, or any other programmable communication device.

[0045] Servers: A computer node on the communication network. Servers are responsible for routing Messages sent back and forth between Users over the network
10 and the persistent storage and archiving of Media payloads. Servers provide routing, transcoding, security, encryption and authentication and the optimization of the transmission of data across the network.

B.2. Media

[0046] Message: An individual unit of communication from one User to another.
15 Each Message consists of some sort of Media, such as voice or video. Each Message is assigned certain attributes, including: (i) the User sending the message; (ii) the Conversation it belongs to; (iii) an optional or user created Importance Tag; (iv) a time stamp; and (v) the Media payload.

[0047] Media: Audio, video, text, position, sensor readings such as temperature, or
20 other data.

[0048] Conversation: A thread of Messages (identified, persistently stored, grouped, and prioritized) between two or more Users on their Devices. Users generally participate in a Conversation using their Devices by either Reviewing Messages in real time or in a time-shifted mode, or creating and sending Messages of a
25 Conversation as desired. When new Messages are created, they either define a new Conversation, or they are added to an existing Conversation.

[0049] Head of a Conversation: The most recent Message of a conversation that has been encoded by the most recent speaker. It is where a User is positioned in a Conversation when reviewing “live” or where one jumps to if the “Jump To Live”
30 feature is used.

[0050] Multiple Conversation Management System or MCMS: An application that runs as part of a Client application, which enables a User to engage in multiple Conversations using a variety of Media types. With the MCMS application, a User

selects one Conversation among the multiple Conversations as current, where only the Messages of current conversation are rendered. For the selected current Conversation, the User may transition from a series of back and forth Messages in time-shifted mode to near-synchronous “live” mode, similar to a standard telephone conversation, and back again. The Messages of the non-selected Conversations are in a paused state. Messages associated with the non-selected Conversion will accumulate if others are still participating in those Conversations. The User may selectively transition the current Conversation among the multiple Conversations and Review the accumulated Messages of the selected current Conversation.

10 **[0051] Multiple Conversation Management System-Consecutive or MCMS-C:** Similar to MCMS, with the added feature of rendering and enabling Users to manage and participate in multiple Conversations consecutively through a hierarchical system of Priorities and time-shifting, which are automatically managed by the system. The MCMS-C application allows the Messages of consecutive Conversations to be rendered in a prioritized order, as opposed to MCMS where only the Messages of the currently selected Conversation are rendered. MCMS-C is particularly applicable in situations where it is important that the Messages of the consecutive Conversations are rendered, in the prioritized order, and/or the receipt of all Messages, even those belonging to lower priority Conversations, is more important than receiving the Messages in real-time. Examples of situations where MCMS-C may be suitable include, but are not limited to, hospitals, taxi fleet management, or tactical communications.

25 **[0052] Multiple Conversation Management System-Simultaneous or MCMS-S:** Similar to MCMS, with the added feature of enabling With MCMS-S, multiple Conversations are selected for simultaneous rendering, as opposed to MCMS where the Messages of only the selected current Conversation are rendered. The MCMS-S application is particularly applicable in situations where a User is listening to multiple Conversations at the same time, such as a trader listening to multiple brokers on different exchanges and periodically sending trading requests to one or multiple of them simultaneously. MCMS-S may also be suitable for tactical communications as well.

30 **[0053] Priority:** The mechanism through which the system determines which Message to render next when a User is participating in MCMS-C. Priority is

automatically managed by the system. A User can set default Priorities, or a predetermined set of system Priorities may be used. In the event of a conflict, where more than one Message is ready to be rendered at the same time, the system resolves the conflict at least partly based on Priority, to determine what Message to render immediately and what Message to time shift.

[0054] Tags: a set of attributes a User or the system may assign to a Conversation or a message, such as a topic (a company name), a directive (“action items”), a indicator (“conversation summary”), or any other label by which one might want to search or organize the data.

[0055] Importance Tags: A special Message attribute that enables a sender to specify when a Message is to be rendered, regardless of other Priority settings. An “urgent” Importance tag will override other Priorities for example. This feature is critical for tactical systems, though any system can be configured to use or disable this feature.

[0056] Packet: Any unit of binary data capable of being routed through a computer network. Each packet consists of header (meta data) and payload (media data). Includes standard packet protocols such as, but not limited to, Internet Protocol (IP), EvDO, UMTS or any other packet-based network, either radio, fiber optic, or wired.

[0057] Header or Packet Header: The portion of a packet that describes the packet; the meta data concerning the payload, its encoding type and destination.

[0058] Vox packet: A proprietary packet that enables the system and method to further refine and optimize the delivery of Messages, Media and other signaling information.

[0059] Media Payload (or Payload): The actual Media portion of a Packet.

B.3. Media Management

[0060] Time Shift Delay (TSD): The amount of time between the arrival of a Vox Packet and the rendering of the Packet on a Device. The TSD must exceed the **Minimum Time Shift Delay**. The TSD is typically determined by the User’s behavior in choosing to review the Messages of a Conversation some time after receipt.

[0061] Minimum Time Shift Delay (MTSD): The time shift delay enforced by the Client to allow for jitter processing using jitter buffer techniques. This causes the system to delay rendering until an adequate number of the packets have arrived to

create a usable media stream. The system will typically adaptively adjust the MTSD over time to compensate for variable conditions in the network.

[0062] Rendering: Delivering a Media stream to a User in a form suitable for User consumption (e.g., voice, text, graphic display, video, or a combination thereof).

5 **[0063] Mixing:** The Rendering of one or more Media streams. For example, the Media stream from two Participants of a Conversation may be Mixed when Rendered, creating a User experience similar to a conversation where multiple people are speaking at the same time.

[0064] Encoding: The process of translating Media either created by a User (such as
10 voice or video) or otherwise originating on a Device (such as GPS or other sensor data), and converting the media into digital data to be processed by a Client.

[0065] Adaptive Jitter Buffer: Jitter buffers or de-jitter buffers are used to counter jitter (i.e., either the arrival of out of sequence packets or the delayed arrival of packets) introduced by packet switched networks, so that the continuous rendering of
15 audio (or video) signals transmitted over a network can be performed without disruption. The data is stored in a buffer before Rendering to allow a reasonably sized buffer of Media to arrive. The Media may be rendered before all the Packets are received, trading off quality for currency. An Adaptive Jitter Buffer is capable of dynamically changing its size to optimize the delay/quality tradeoff.

20 **[0066] Persistent Infinite Message Buffer (PIMB):** The PIMB is a storage management system for the storage of time-based Media that performs both the de-jittering of “live” data and the storage and retrieval of archive data. The PIMB further includes the additional attributes of potentially infinite and persistence storage of Media. The PIMB maintains “exact” or full copies of Vox Packets of a Message and
25 Conversations at some or all Participant Devices and/or Servers.

[0067] Packet Loss Compensation or Concealment) (PLC): During Rendering of a Media stream, the PLC component compensates for missing Packets, interpolating the results to present the stream to a reviewer. Missing Packets may be rendered as silence, or information from adjacent Packets may be used to present an interpolated
30 sound or image. The particular method to be used will be dependent on the media, Codecs in use, and other generally known parameters.

B.4. People

[0068] User: A person who is authorized to use the system.

[0069] **Contact:** A record of either a User or non-user of the system. Users typically engage in Conversations with members on their list of Contacts. A non-user is a user that accesses or uses the system using a legacy phone, radio or other non-Client 12 enabled device.

5 [0070] **Group:** The association of multiple Contacts. Contacts may be selectively added or deleted from a Group. When a Conversation takes place among a Group, all the members of the Group may or may not participate.

[0071] **Channel:** Typically used for tactical communication systems. A Channel is similar to a Group in that it associates multiple Contacts with the Channel.

10 [0072] **Participant:** A person who is identified as a member of a Conversation. Could be a User or a non-User participant.

B.5. Conversation Management

[0073] **Time Shifting:** Time shifting is the ability to play any Message at any time after it has been received as determined by the User-recipient. By Time-Shifting, a
15 User may Review a Message: (i) immediately on demand by Rendering immediately after the MTSD; or (ii) time-shifted in a mode of reviewing the Message upon the discretion of the User; (iii) from the archive for searching, reconstructions, etc. of old Conversations; (iv) after a delayed period of time to accommodate the Reviewing of other higher Priority Messages (or Conversations) that need to reviewed first; (v)
20 and/or repeatedly if necessary for the Message to be reheard and understood. In other words, Time Shifting is the ability of a user to render a Message at any time after the system imposed MTSD.

[0074] **Reviewing:** Listening, viewing, reading or otherwise observing the Media content in Messages. Reviewing may take place in either a near synchronous real-time
25 “live mode” or the time-shifted mode.

[0075] **Intention:** Either (i) a User-defined attribute capturing whether the User wants to Review the Messages of a Conversation either as soon as possible or Review the Messages in a time-shifted mode; (ii) implied by a User’s behavior; or a combination of (i) and (ii). .

30 [0076] **Attention:** A user attribute capturing whether the User is Reviewing the Messages of a given Conversation at the moment.

[0077] **Catch Up To Live (CTL):** A rendering mode that allows a User, who is not at the Head of a Conversation, to Review previous Messages more quickly to “Catch Up

To Live” (i.e., the Head of the Conversation). The CTL feature may use any of a number of catch up techniques, such as the faster replay of Messages, the removal of gaps in the Media of the Messages, removal of hesitation particles, etc. When the User has caught up to live, the system seamlessly flows into the live Conversation.

5 This is a very useful feature with conference calls, for example, in situations where a User needs to temporarily shift their attention away from the Conversation, but wishes to hear the entire Conversation upon their return.

[0078] Catch Up Mode: A user-configured or pre-configured mode that determines how the CTL process will catch-up (i.e., replay faster, remove silence, and hesitation
10 particles, or a combination thereof).

[0079] Jump To Live (JTL): This feature allows a user to jump from their current position to the Head of a Conversation. A user will typically use the JTL feature when they do not want to Review all of the Messages between their current position in the Conversation and the Head. When the JTL feature is implemented, the user skips over
15 any intervening Messages and starts Rendering the “live” Messages at the head of the Conversation.

[0080] MCMS Participant Attributes: A set of attributes, either defined by a User, interpreted by the system from the User's behaviors, assigned by an administrator, or a combination thereof, which define the Intention, Attention, Priority, and rendering
20 preferences of a receiver for a given Conversation. The attributes include, but are not limited to: (i) the Intention of when a receiver would like to render to the Messages of the Conversation. Possible Intention values include: “now”, “time-shifted”, “Catch Up To Live” (CTL), “paused”, and “never”; (ii) Catch Up Mode, which is a configuration setting which determines how the CTL process should catch the receiver up to live
25 (e.g., play faster, skip silence gaps or hesitations, or play at normal speed); (iii) Time Shift Delay (TSD), which defines how far the receiver's current position in the conversation is from the Head of the Conversation, and (iv) the Priority of the Message with regard to the receiver's other Conversations.

C. System Architecture

30 **[0081]** Referring to Figure 1, a block diagram of the telecommunication and media management system according to one embodiment of the invention is shown. The system 10 includes a plurality of Clients 12₁ through 12_n, running on Devices 13₁ through 13_n respectively. The Devices 13 communicate with one another over a

communication services network 14, including one or more Servers 16. One or more networks 18₁ through 18_n, is provided to couple the plurality of Devices 13₁ through 13_n to the communication services network 14. In various embodiments, the networks 18 may be the Public Switched Telephone Network (PSTN), a cellular network based on CDMA or GSM for example, the Internet, a tactical radio network, or any other communication network, or a combination thereof. The communication services network 14 is a network layer on top of or otherwise in communication with the various networks 18₁ through 18_n. In various embodiments, the network layer 14 is either heterogeneous or homogeneous. Clients 12₁ through 12_n communicate with one another and with Servers 16 over the networks 18₁ through 18_n and network 14 using individual message units referred to as “Vox packets”, which are described in detail below.

D. Client Architecture

[0082] Referring to Figures 2A and 2B, a block diagram of a Client 12 running on a Device 13 is illustrated. As illustrated in Figure 2A, the Client 12 includes Multiple Conversation Management System (MCMS) application 20, a rendering and encoding module 21, and an MCMS applications database 22. As illustrated in Figure 2B, the Client 12 further includes a Store and Stream (SAS) module 24 with a Persistent Infinite Message Buffer (PIMB) reader 26, a PIMB writer 28, PIMB database 30, a data and network quality (DNQS) store 32, and Media driver and encoder hardware 34. The MCMS application 20 and the Store and Stream module 24 communicate with one another through message handling modules 25a and 25b respectively. The Client 12 further includes an authentication-encryption-security module 40 and a communication protocol module 44.

[0083] The module 40 provides authentication, encryption and security services during the transmission and receipt of “Vox” packets to and from the Client 12. The communication protocol module 44 encapsulates Vox packets into the native packets used by the underlying network 18 connected to the Device 13 running the Client 12 when transmitting data and de-encapsulating Vox packets from the native packets when receiving data. With the modules 40 and 44, multi-party end-to-end authentication, encryption and security is provided between Clients 12. Messages are authenticated, encrypted and secured across the networks 18₁ through 18_n and network 14, from a first sending Device 13 to second receiving Device 13.

D.1.1. The MCMS Database

[0084] The database 22 stores and manages the persistent meta data for a number of entities in the system 10, including Contacts and Participants, Conversations and Messages (live and stored), and default Priorities, and information regarding the Servers 16. In addition, the MCMS database 22 stores the moment-to-moment operational data of a User's Conversations, presence, and status, as well as that of all the Participants conversing with the User or on the User's Contact list. For example, with regard to Conversations and Messages, the database 22 keeps track of status information, such as what Messages of a Conversation a User has or has not Reviewed, Priorities, and Catch Up To Live status for each Conversation in which the Client 12 is a Participant, the presence and status of all Participants, and other network and other system management data.

D.1.2. The MCMS Application

[0085] MCMS application 20 supports the different Voicing modes of engaging in conversations and/or managing multiple conversations using a variety of media and data types (voice, video, text, location, data, etc.). Users engage in Conversations by sending Messages to designated recipients using their Client 12 enabled Devices 13. Depending on preferences and Priorities, the recipient might Review the Message in real time, or they might simply be notified that the Message is ready for Reviewing. Users can transition from a series of back and forth Messages, which are Reviewed in a time-shifted (or voice messaging) mode or in a near synchronous, full duplex conversation (similar to standard "live" phone calls) and then back to voice messaging again. The MCMS application 20 allows a User to control their interactions with their most important Conversations in real-time without missing any Messages in other ongoing Conversations. For example, the MCMS application 20 notifies a User of urgent or high priority communications from a Conversation that they are not currently Reviewing. MCMS application 20 also enables all Messages from all Conversations to be stored for later retrieval so they can be reviewed at any time.

[0086] In accordance with various embodiments, there are several different operational modes of the MCMS application 20, including MCMS-Consecutive (MCMS-C) and MCMS-Simultaneous (MCMS-S), which support the consecutive and simultaneous rendering of Messages respectively. Each of these embodiments is described in more detail below. Unless specifically specified, the term "MCMS" is

intended to generally mean the MCMS application 20, which includes the aforementioned different modes.

[0087] The MCMS application 20 is a multi-tiered architecture that includes a number of modules and services. The modules and services include the MCMS Database Module 20a, the SAS Services Module 20b, the Messaging and Signaling Services Module 20c, the User Interface Application Programming Interface (API) 20d, the User Interface Module 20e, the Conversations/Messages Management Services 20f, the Priorities Services 20g, the Contacts Service 20h, the Presence/Status Services 20i, and the Messages/Signals Services 20j.

10 **D.1.2.1 The MCMS Database Module**

[0088] The MCMS database module 20a is a service module that manages all function calls necessary for the MCMS application 20 to access the MCMS database 22.

D.1.2.2 The SAS Services Module

15 [0089] The SAS Services module 20b includes a set of function calls that enable communication and coordination between the MCMS application 20 and the Store and Stream module 24, and which are passed back and forth through the message handling modules 25a and 25b respectively. The set of function calls enable both the MCMS application 20 and the Store and Stream module 24 to operate as necessary to
20 implement the various Voxing functions when invoked by Users and/or as dictated by network conditions. Some of the functionality performed by the SAS Services module 20b includes maintaining and communicating the status of Message transmissions and Message acknowledgments, the instructions for rendering of Messages, and the status and presence of Users.

25 **D.1.2.3 The Messaging and Signaling Services Module**

[0090] The Messaging and Signaling Services module 20c runs on both Clients 12 and Servers 16 and enables communication between the Client 12 and the Servers 16 of the system 10. This communication, which includes messages, data and other signals, allows the Client 12 and the system 10 to track and administer
30 communications, network status, Users, and User status. Types of messages and signals sent between the Message and Signaling Services modules 20c running on the Clients 12 and the Servers 16 include, for example, network availability of Users, tracking of Messages that the Server 16 has sent to the Client 12 (possibly including a

“high water mark”) to determine if an entire message or some portion of a message is missing, (e.g., a sequence number per Participant per Conversation created by the “generating” Client), whether a user is speaking or Reviewing Messages of a given Conversation, where a User is with respect to the Head of a Conversation, or when a Participant is no longer Reviewing a Conversation live. These are examples a few of the many types of messages and signals sent between the Message and Signaling Services modules on the Clients 12 and Servers 16 and in no way should be construed as limiting the invention.

D.1.2.4 The User Interface API

10 [0091] The User Interface API 20d is a module that defines a set of function calls that define the programming interface between the User Interface module 20e and the underlying services of the MCMS application 20. The User Interface API 20d supports general-purpose methods such as UI application support, and all function calls necessary for a User Interface to operate the MCMS Application 20. In various
15 embodiments, the User Interface API 20d enables the Client 12 to support a wide variety of user interfaces and device types, such as Adobe Flash-based and/or Microsoft Windows applications, cellular or mobile phone devices, PSTN devices driven with tones, a voice user interface (VUI), and physical radio communication interfaces. In various embodiments, the User Interface API module 20d enables the
20 design of both highly flexible and highly constrained user interfaces to support the functionality of the MCMS application 20.

D.1.2.5 The MCMS User Interface Module

[0092] The MCMS User Interface module 20e supports the operation and functions of the audio and video user interface of the Client 12. The User Interface module 20e
25 supports a host of user interactions and can be implemented with various interaction mediums, such as, an array of graphical user interface screens, an Audio/DTMF interface, or voice user interface on the Device 13, all of which enable a User to interface with the system 10. A partial list of User interactions that are supported include, for example, functions to: log-in; manage, join, and monitor Conversations;
30 control Conversation rendering; manage Priorities; and requests to review archived Conversations. It should be noted that this list is exemplary and in no way should be construed as limiting the invention.

D.1.2.6 Conversation/Message Management Services

[0093] The Conversation/Message management services 20f is a module which defines a set of functions that manage the data structures and processes responsible for managing and retaining all information needed for the User to manage the receipt and Review of transmitted and received Media (e.g., voice or video content
5 Messages) between the participants of a Conversation. The Messages are organized into Conversations. Media that is sent or received by the Device 13 running the application 12 is available for immediate Review while being received. The received Media is also recorded for Review in a time-shifted mode, Conversation management,
10 and archival purposes. In an alternative embodiment, Messages or Conversations can be optionally marked for transience, specifying their desired retention requirements (e.g., some Messages will not be retained or stored beyond the requirements for immediate rendering). In yet another embodiment, Media can be optionally marked for review in a time-shifted mode only and cannot be reviewed immediately upon
15 receipt.

[0094] The Conversation/Message management services 20f further enables, for each current or ongoing Conversation of the User, the sending of Media to a receiving Client 12 at any time, and the receiving Client 12 seamlessly associates these Messages with the appropriate Conversation, regardless of the actions or inaction of
20 the receiver.

[0095] With the Conversation/Message management services 20f, all Conversations are essentially asynchronous. If two Users are actively engaged in a given Conversation and the User controlled delay between transmissions is minimal, the experience will be one of a synchronous full duplex conversation, as with current
25 telephone or VoIP conversations. If either User delays their participation, for whatever reason, the Conversation drifts towards an asynchronous voice (or other Media) messaging experience. In alternative embodiments, Conversations can be optionally Tagged as asynchronous Messages only or synchronous Messages only. In either of these cases, the Conversation cannot drift between the two modes, unless the
30 Tag is reset. After the Tag is reset, the Conversation again may flow between near synchronous (i.e. live or real-time) and asynchronous (i.e., time-shifted or voice messaging) modes.

[0096] The Conversation/Message management service 20f processes the transmission and receipt of Messages in a progressive fashion. When transmitting, Media may be created while Messages are simultaneously encoded, stored and transmitted. In other words, the transmission of Messages may occur simultaneously with the generation of Media by the User (i.e., while speaking into their Device 13 or generating video). On the receiving side, the receipt, storage, and Rendering of Messages also all occur progressively. Messages do not need to be completely received before they can be Rendered. The Rendering of Messages may occur at the same time Messages are being delivered, right up to the MTSD. Further, the service 20f is also capable of the simultaneous transmission of outgoing Messages and Rendering of incoming Messages. The progressive nature of the service 20f allows a User to be engaged in a live Conversation while storing and streaming the media of the Conversation for later retrieval and review as well other functions described herein.

[0097] The time-shifting of Messages by the Conversation/Message management services 20f allows a User to “Catch Up To Live” on a Conversation if they missed earlier Messages or were involved in another Conversation. This time-shifting process eliminates the need for Users to broadcast a request to their entire Group or Channel to have Messages repeated. Older Messages may be replayed at any time at potentially higher speeds to save time. Users may easily skip forward and backward through their Messages and within individual Messages. The Reviewing process may be configured on a Message-Priority basis to potentially skip lower priority Messages.

[0098] In one embodiment, the Conversation/Message management service 20f also identifies Messages by a specific Participant (speaker) and, by default, mixes Messages of a Conversation that were delivered at the same time (MCMS-S). In an optional embodiment, a User could Review the transmissions of different Participant speakers of a Conversation separately.

[0099] The Conversation/Message management module 20f further allows Conversation sharing among Participants, who can be added to an active or archived Conversation. In one embodiment, an added Participant to a Conversation is provided access to the Conversation and has the ability to retrieve the previous Messages of the Conversation for Review. In an alternative embodiment, the added Participant is

provided access to the Messages of the Conversation only from the point where the new Participant joined, and not any previous Messages of the Conversation.

[0100] The Conversation/Message management module 20f is also responsible for managing the functions used to control all rendering tasks performed by the Store and Stream module 24. These tasks include rendering Media (i.e., voice, video, etc.) appropriately for the Device 13 running application 12. These rendering tasks include, but are not limited to, the rendering of Mixed Messages (i.e., overlapping messages), as well as rendering according to user-defined criteria, such as playing faster, catching up to live, removing silence, removing hesitation particles, frequency shifting, and the ability to apply independent gain control to individual senders in a multi-party conversation.

D.1.2.7 Priority Services

[0101] The Priority service 20g is a module that defines a set of functions that manage the data structures and processes responsible for managing and retaining all information needed for a User to manage the Priority of the consecutive Conversations (i.e., MCMS-C) in which the User is engaged. When a User participates in a number of consecutive live Conversations, the User is required to prioritize the Conversations. Issues arise when Messages of different Conversations are ready to be rendered at the same time. An algorithm is used to determine the order in which the Messages are rendered that considers the availability of Messages to be rendered and the Priorities set by the User. The algorithm determines that the available Messages with the highest priority are rendered first while any concurrently available Messages are time shifted automatically just enough to allow for the rendering of the higher priority Message. As rendering time becomes available, the system will automatically render the time-shifted messages according to the User's Priorities.

D.1.2.8 Contacts Services

[0102] The Contacts services 20h is a module that defines a set of functions that manage the data structures and processes responsible for managing and retaining all information needed for authenticating and associating one or more Contacts with Conversations. When sending a Message as part of a Conversation that is associated with a number of Contacts, all of the Contacts receive the Message.

D.1.2.9 Presence/Status Services

[0103] The Presence/Status services 20i is a module that defines a set of functions that maintain the data structures and processes responsible for managing and sharing presence and status information between certain Users and/or non-users of the system.

5 In various embodiments, the presence and status information is maintained for all User and non-users engaged in the Conversations the User of the Client 12 is engaged in, all Users and non-users in the Contacts list, or Users within a predefined domain (e.g., the members of a corporation or other organization). These examples are merely illustrative and should not be construed as limiting. The Presence/Status services 20i

10 module may manage and share presence and status information on any defined set of Users and/or non-users.

[0104] The Presence/Status service 20i enables Users to monitor the status of other User's Intentions, Attention, and their Time Shift delay on any given Conversation (i.e., how far back they are from Reviewing the Messages of the Conversation live).

15 In one embodiment, privacy controls are provided concerning availability of Presence and Status data. The Presence/Status module 20i further controls the data that enables the system 10 to deliver Messages that match the behavior and Intentions of Users. For example, a User may indicate their Status by designating an Intention to either Review or not Review a Conversation live. In response, the Presence/Status services

20 20i issues commands that cause the rendering of Messages either "live" or time-shifted, in accordance with the Intention of the User. In addition, the Intentions of Users are shared with the other Participants of the Conversation. The service 20i is also capable of inferring other Status values from the User's behaviors. The Presence and Status information is also used to optimize network traffic and bandwidth, as

25 described in more detail below.

D.1.2.10 Messages/Signals Services

[0105] The Messages/Signals services 20j is a module that defines a set of functions that manage the data structures and processes responsible for messaging and signaling Users of the system 10, using special messages or audible tones. The special messages

30 or tones may include for example an indication if a Message or Messages are live or time-shifted, whom the Message(s) are from, Priority, and other factors. The Message/Signal service 20j further has the ability to (i) signal the presence or absence of Users on the network, as well as the ability to notify if one or more Users are no

longer actively Reviewing the Messages of a Conversation; (ii) “ring” or otherwise notify another User to get their attention when they are paying attention to another Conversation or not paying attention to their Device 13 at all; (iii) leave a Message for Users currently not on the network 18 to immediately review the Message the next
5 time the individual re-connects to the network 18; (iv) generate an audible or visible feedback that alerts the sender that a sent Message was not received by recipient(s), generate a confirmation when the Message has been received by the recipient(s), and/or a confirmation indicating when the Message has been Listened to by the recipient(s); and (v) implement a Priority scheme where individuals on a Conference
10 or tactical call may be notified that their attention is immediately needed on the call. This indication may convey multiple levels of urgency and an acknowledgement of some kind by the recipient.

D.1.2.11 Rendering and Encoding

[0106] The rendering and encoding module 21 is responsible for performing all
15 rendering tasks for the MCMS application 20. These tasks include rendering Media appropriately for the device 13 running application 12.

D.2 The Store and Stream Module

[0107] The Store and Stream module 24 supports a number of functions and performance attributes, which are described below.

20 [0108] With the Store and Stream module 24, Message transmission is essentially “full-duplex”, enabling any party to send a Message at any time, even while another party is also sending a Message, or if the other party is unavailable or otherwise engaged. The Store and Stream module is able to render messages as in a live PSTN or VoIP call or deliver them for time shifted messaging modes. It is able to optimize
25 transmission and control Rendering according to the desires of the User.

[0109] The Store and Stream module 24 maintains connectivity with all target recipients (e.g., Servers 16 or other Devices 13) on the underlying network 18, manages all message, signal, and media transmissions, and optimizes the delivery speed and bandwidth usage across the network 18 to meet a User’s immediate
30 performance requirements, while managing network quality and capacity. The module 24 adapts and optimizes Media delivery commensurate with the quality and capacity of the underlying network 18. When insufficient underlying network resources are available, the quality of the transmitted Media streams can be degraded. As bandwidth

becomes available, the quality of the transmitted Media streams may be increased. In addition to tradeoffs of Media quality, the Store and Stream functionality can make tradeoffs in the amount of Media transmitted in each packet based on Users' intentions to render data in real time as described below.

5 [0110] By dynamically controlling the delivery rate of Media based on the conditions of the underlying network 18, the Store and Stream module 24 is optimized to deliver time-sensitive Media that is "good enough" to Render upon receipt, and the guarantee eventual delivery of exact or full copies of the Media for archival purposes through a background process of requesting retransmission of missing, low quality, or damaged
10 packets. As long as sufficient network resources exist to meet minimum Media quality levels, this retransmission does not impede the Rendering of live call Media. The Clients 12 of the system 10 are thus designed to bridge the performance gap between the delivery of an exact or complete copy of the Media at the expense of substantial potential latency versus the quick delivery of Media, but with no
15 guarantees of completeness. In the context of this application, the term "good enough" means that the quality of the Media is sufficient so that when it is rendered, it is intelligible. The notion of "good enough" is therefore subjective and should not be construed in absolute terms. For example, the quality level of certain Media to be good enough may vary depending on the type of Media, circumstances, and other
20 factors.

[0111] The Store and Stream module 24 further persistently stores all Media created by or otherwise originating using a Device 13 or received over the network 18 from other Device 13 and/or users. There are several significant advantages of storing this Media on the Device 13 running the Client 12: (i) it enables Users to leave a Message
25 for another party, even when the sender and/or the recipient has either unavailable or poor network connectivity. In the case of insufficient bandwidth, the Message will be transmitted as fast as available bandwidth can be effectively used. In the case of no connectivity, the Message is queued for transmission as soon as network connectivity becomes available, resulting in a time-shifted delivery; (ii) the User has the ability to
30 pause, replay, fast-forward, and Catch-Up-To-Live with an ongoing Conversation, as well as retrieve and review the archived Messages of previous Conversations; and (iii) it enables the optimization of data payloads over the system 10 and improves system

resilience against network bandwidth and connectivity problems that may occur from time to time.

[0112] The Store and Stream module 24 is also responsible for: Mixing Messages as appropriate to create overlapping Messages (generated by the normal overlap of speakers in a Conversation or background noise), simulating an actual Conversation where multiple parties are speaking; rendering transcriptions or translations of audio media; adjusting the rendering of Media according to a number of User-defined criteria including playing faster, removing silence gaps between spoken words, removing hesitation particles, and frequency shifting; and the ability to apply independent gain control to individual senders in a multi-party Conversation; as well as other potential Rendering options.

[0113] The Store and Stream module 24 manages control and informational messaging between itself and MCMS.

D.2.1 The Persistent Infinite Message Buffer (PIMB)

[0114] The Persistent Infinite Message Buffer or PIMB 30 is a set of indexed (i.e., time-stamped and sequentially numbered) Media payload data structures and a system for their storage and retrieval. In one embodiment, the data in the PIMB 30 is arbitrarily persistent, meaning it is available virtually forever or at least until the system runs out of storage. Various retention rates and strategies may be employed to make effective use of storage resources. Many possible implementations exist for the physical storage implementation of the PIMB 30, including, but not limited to: RAM, Flash memory, hard drives, optical media, or some combination thereof. The PIMB 30 is also “infinite” in size, meaning the amount of data that can be stored in the PIMB 30 is not inherently limited. This lack of limit is in comparison to existing jitter buffer technology that discards data as soon as it is rendered. In one specific embodiment, the PIMB 30 may be implemented using a small and relatively fast RAM cache memory coupled with a hard drive for persistent storage. As the physical storage capacity of the PIMB 30 is exceeded, the data is maintained on the Server 16 (as described below) for later retrieval on demand. User criteria or a replacement algorithm, such as least-recently-used, or first-in-last-out, is used to control the actual data stored in the PIMB 30 and the data that is maintained on the Server 16 or archived at any point in time. The PIMB 30 further provides the attributes of file system storage and the random access attributes of a database. Any number of

Conversations, regardless of their duration or the number of Messages in each, may be stored and later retrieved for Review. In addition, the meta data associated with the Messages of a Conversation, such as its originator and its length, may be also stored in the PIMB 30. In alternative embodiments, the indexed Media payloads and other data can be stored for a designated period of time (e.g. 30 days). Once the age of the media exceeds the designated period, the payloads and data are discarded. In another embodiment, payloads can be discarded based on the sender and/or the recipient of the Message containing the payload, or the topic of the Conversation or Messages associated with the payload. In yet other embodiments, payloads and data may be marked for transience, meaning the Messages will not be stored in the PIMB 30 beyond the requirements for immediate Rendering.

D.2.2 The Data and Network Quality Store

[0115] The data and network quality store (DNQS) 32 is a data store for storing information regarding Media payloads and Vox packets that are either read from or written to the PIMB 30.

D.2.3 The PIMB Writer

[0116] The PIMB writer 28 writes data to the PIMB 30 for two basic purposes. The PIMB writer 28 writes data from a Media capturing device (e.g., a microphone or camera) on the Device 13 running the Client 12 (“Encode Receive”). The PIMB writer 28 also writes data received over the network 18 from other Clients 12 into the PIMB 30 (“Net Receive”).

D.2.3.1 Encode Receive

[0117] For capturing Media from the Device 13, the PIMB writer 28 includes Encoder Receiver 28a and a Data Storer 28c. When a User speaks into the microphone or generates video images with the Device 13 for example, the hardware 34 receives the raw audio and/or video signals and provides them to the Encoder Receiver 28a, which encodes the signals into indexed media payloads (hereafter sometimes simply referred to as “payload” or “payloads”). The Data Store 28c stores the payloads into the PIMB 30. Other types of Media, such as sensor data is converted into payloads in a similar manner.

D.2.3.2 Net Receive

[0118] For storing Media received over the network 18 into the PIMB 30, the Net Receive function of PIMB writer 28 includes a Network Receiver 28d, a Data

Bufferer 28e, a Data Storer 28f, a Data Quality Manager 28g, and a Network Quality Manager 28h. The Network Receiver 28d receives Vox packets over the network 18. The Data Bufferer 28e places the received Vox packets into their proper sequence and prevents the Rendering of the incoming Vox packets for at least the Minimum Time Shift Delay (MTSD) delay. The Data Storer 28f transforms the packets into indexed media payloads and stores the indexed media payloads in the PIMB 30. As the payloads are stored, the Data Quality Manager (DQM) 28g notes any missing or defective packets. If packets are missing or defective, the DQM 28g schedules a request for retransmission over the network 18. The sending device replies by resending the missing or defective packets. Eventually these packets are converted into indexed media payloads and stored in the PIMB 30. By retrieving the missing or defective packets, an “exact” copy of a sender’s Message is eventually stored in the PIMB 30. The retransmission of missing and/or defective packets does not delay the rendering of Messages in real time, provided that the packets that have been delivered are of “good enough” quality and quantity. Retransmission requests may be delayed by the DQM 28g if insufficient network resources are available to support both the new “live” data as well as the retransmission.

D.2.4 The PIMB Reader

[0119] The PIMB reader 26 reads data from the PIMB 30 for two basic purposes. The PIMB reader 26 accesses the PIMB 30 when data is to be rendered (“Render”) for the local Client 12. Data is also read from the PIMB 30 when data is to be transmitted (“Transmit”) by the Client 12 over the network 18.

D.2.4.1 Render

[0120] For the rendering of messages on the Client 12, the PIMB reader 26 includes a Data Prioritizer 26a, a Data Retriever 26b, a Packet Loss Compensation/Interpolator (“PLC/Interpolator”) 26c, a Data Mixer 26d and a Data Renderer 26e. The Prioritizer 26a prioritizes the data to be rendered by building an ordered queue of messages that could potentially be Rendered. It uses User configured Priority for the rendering of consecutive Conversations (MCMS-C). In addition, the Data Prioritizer uses the availability of media data to render within the limits imposed by the MTSD, the User’s current Attention, and the User’s defined and implied Intentions. The Data Retriever 26b retrieves the prioritized indexed media payloads from the PIMB 30. The PLC/Interpolator 26c performs packet loss compensation and interpolation on the

retrieved payloads, using known packet loss compensation and interpolation algorithms. The particular method to be used is dependent on the media Codecs in use, and other well-known parameters. The Mixer 26d is used to appropriately mix concurrent data streams from multiple Messages of a single Conversation together.

5 For example, if two or more Participants of a Conversation are speaking at the same time, the Mixer 26d mixes the Messages, creating the effect of both Participants speaking at the same time. In an alternative embodiment, the User has the option of Reviewing the multiple streams from one Participant at a time. If only one Participant in the Conversation is speaking, the Mixer 26d may simply pass the single Message

10 stream, without performing any mixing. The Renderer 26e takes the data from the Mixer module 26d and converts it into a form suitable for the hardware driver 34. The hardware 34 then drives either the speaker or video display of the Device 13, depending on the type of Media, creating voice, video or some other audible and/or visible notifier on the Device 13.

15 **D.2.4.2 Transmit**

[0121] To prepare messages for transmission from the Client 12 over a network 18, the PIMB reader 26 includes a Data Prioritizer 26f, a Packet Quality Manager (PQM) 26g, a Data Retriever 26h, Packetizer 26i, a Transmitter 26j and an Acknowledger 26k. The Data Prioritizer 26f prioritizes the Messages for transmission over the

20 network 18. Priority is determined using the MCMS Participant Attributes related to payloads available for transmission, network connectivity and bandwidth conditions, and the Intentions of Users beyond the next hop to either Render live or time-shifted, and, in some embodiments, possible optimizations of transmission bundling where multiple packets to any given next network hop are available. The prioritized packets

25 are then optimized, using the PQM 26g, which assures the timely delivery of “good enough” data quality for live Messages, while minimizing real-time bandwidth, as described in more detail below. The Data Retriever 26h retrieves the appropriate payloads from the PIMB 30. The Packetizer 26i assembles the payloads into Vox Packets, which are then transmitted by the Transmitter module 26j over the network

30 18. When the recipient receives the Vox packets, an acknowledgement is sent back to Acknowledger 26k over the network 18 for notifying the sending User that a Message has arrived at its destination.

D.2.5 The Packet Quality Manager

[0122] The PQM 26g has several optimization goals: (i) the timely delivery of an adequate copy of time-sensitive Media (“as soon as possible to be good enough” for Rendering); (ii) the efficient use of available bandwidth, meaning using the optimal transmission frequency, payload quality, and packet size for the underlying network; and (iii) the ability to dynamically adjust or make changes in the transmission frequency, payload content, payload quality, packet size, etc. as network conditions change.

D.2.6 The Network Quality Manager

[0123] On the receiving side of a network transmission is the Network Quality Manager 28h (NQM). The NQM is responsible for observing specific properties of network performance for each sender that has sent media to the Client 12, comparing expectations of jitter, loss, and throughput to their actual values. This is used to compute a Network Quality Rating (NQR) for every sender. This NQR is used to indicate sender availability and Conversation live-ness to the User of the receiving device.

D.2.7 The Data Quality Manager

[0124] The Data Quality Manager 28g measures the quality of the data being received over the network by observing packet loss, jitter, and throughput. The DQM 28g uses these measurements for three purposes: (i) to send receipt reports back to the sender; (ii) optionally using those receipt reports to request retransmission of certain data; and (iii) making these measurements available to the NQM 28h.

E. Server Architecture

[0125] Referring to Figure 3, a block diagram of the application 78 that runs on the Server(s) 16. The application 78 is similar to the Client application 12 in many regards and includes an MCMS server application 80, an MCMS database 82, a store and stream module 84, a PIMB 85, a data and network quality store (DNQS) 86, MCMS-SAS message handling modules 87a and 87b which manage messages and signals back and forth between the MCMS server application 80 and Store and Stream module 84, an archive/retriever 88, and an archive 89. The application 78 further includes an authentication-encryption-security module 40 and a communication protocol module 44.

[0126] The MCMS server application 80 is a multi-tiered architecture including a MCMS Database Module 20a, a Store and Stream (SAS) Module 20b, a Messaging/Signaling Module 20c, Conversations/Messages management services 20f, Priority services 20g, Contacts (including User and Authentication) services 20h, Presence/Status service 20i, and Messages/Signals services 20. The aforementioned modules and services of the application 78 are similar or the same as the modules and services having like reference numerals as the Client 12, and therefore are not described in detail herein, except for one notable exception. In various embodiments, the MCMS server application 80 and Store and Stream module 84, including the MCMS database 82, is configured to support many Users in one instance of the application. The one instance may be further configured to support multiple domains, where each domain is defined as a group of Users (i.e., a corporation or other group of Users belonging to a common organization). This architecture allows each application 78 on Server 16 to serve multiple users (or domains), where each user (or domain) is not visible to other Users. This partitioning is referred to as “multi-tenancy”.

[0127] The server Store and Steam 84 module performs the functions of Net Receive and Transmit. For the Net Receive function, the module 84 includes a Net Receiver 28d, Data Bufferer 28e, a Data Storer 28f, a Data Quality Manager (DQM) 28g, and a Network Quality Manager 28h. For Transmit functions, the module 84 includes a Data Prioritizer 26f, Packet Optimizer 26g, Data Retriever 26h, Packetizer 26i, Transmitter 26j and an Acknowledger 26k. The aforementioned elements of the Store and Stream module 84 are similar or the same elements having like reference numerals as the Client 12, and therefore are not described in detail herein.

[0128] Since the Server 16 does not directly interact with Users, the encoding and rendering functions provided in the Store and Stream module 24 of the Clients 12 need not be present. The MCMS application 80, when running on Servers 16, does not interact directly with Users. Consequently, the user interface and user interface API modules and services 20e and 20d are not needed.

[0129] The application 78 on each Server 16 potentially serves multiple tenants, meaning it serves multiple Users of the system 10. The PIMB 85 of the server application 78 is therefore significantly larger and is used to store the media payloads of multiple Users, as opposed to the PIMB 30, which is used to store just the generated or received payloads of only a single User. The main purpose of the Store

and Stream module 84 is to receive Messages transmitted by the Clients 12 and transmit Messages to other Clients 12. As Messages are received, they are stored in the PIMB 85 and transmitted to the next Server 16 (i.e., the net “hop”) of the network layer 14 along the path to the intended recipient(s), or to the recipient(s) directly depending on the system configuration. The archive-retriever 88 is responsible for archiving the media payloads stored in the PIMB 85 in archive 89. As the physical space in the PIMB 85 is exhausted, media payloads in the PIMB 85 are moved to the archive 89, which is a mass storage device. In various embodiments, the payloads stored in the PIMB 85 may be archived in accordance with User defined criteria and/or any known replacement algorithm, such as first-in-first-out (FIFO) or least recently used (LRU). It should be noted that only a single Server 16 is illustrated in Figure 1 for simplicity. It should be understood that in actual embodiments, multiple servers or a “server farm” may be used for a network with a large number of Users.

[0130] The terms “persistent” and “infinite” used to describe the PIMB 30 and the PIMB 85 should not be construed literally as absolute terms. A User may wish to indefinitely store some Messages that are considered important. In other situations, such as a casual chat between two friends, the Messages may be deleted after a certain period of time to save space. According to various embodiments of the invention, different retention policies may be used, either set by the system 10 or configured by the User. The use of the word “infinite” refers to the lack of any preset time boundaries enforced by the PIMB. This is contrasted with current jitter buffer systems, which discard media after it has been rendered. The terms persistent and infinite should therefore be broadly construed to mean that the PIMB 30 and PIMB 85 have no internal limitations on the time ranges and quantities of Messages that can be stored therein.

[0131] There are a number of advantages to archiving the Messages of Conversations in a persistent storage medium. Voice messages and other Media can be organized, indexed, searched, transcribed, translated, and Reviewed as needed. Voice, as well as other Media, therefore becomes an asset that can be managed both by Users and organizations. These Media assets have value to corporations, first responders, police and fire departments, as well as the military.

F. The Vox Protocol and Indexed Media Payloads

[0132] As noted above, the Vox protocol is used by the Store and Stream module 24 to support all facets of payload transmission, storage, and optimization. The Vox packet is a structured message format designed for encapsulation inside a transport packet or transport packets of the underlying technology of the network 18. This arrangement significantly improves the flexibility of the system 10. By embedding the Vox packets into existing transport packets, as opposed to defining a new transport layer for “Voxing” applications, the system 10 takes advantage of current packet based communication networks running over the existing telecommunications infrastructure. A new network infrastructure for handling the Vox packets therefore need not be created to take advantage of all the benefits of the system and method described herein.

[0133] Referring to Figure 4A, the general format structure of a Vox packet 95 is illustrated. The format of the Vox packet 95 includes fields for type, sub-type, length, and payload. The different types of Vox packets include authentication, signaling, media payload, media multiplex (one message), and media multiplex (multiple messages). The sub-type field is used to designate different types of authentication, signaling or media type messages. Possible sub-types for authentication Messages include those necessary for key exchanges and Authentication. Possible sub-types for signaling Messages include registration, routing, message set-up, and network management. Possible sub-types for Media messages include different Codec styles and different payload aggregation techniques. The length field defines the overall length or size of the payload. The payload field contains the actual payload or media of the packet 95.

[0134] Referring to Figure 4B, a diagram illustrating a Vox packet 95 encapsulated in an exemplary protocol used by the network 18 is shown. In this example, the Vox packet 95 is embedded in underlying UDP, IP and Ethernet transport packets 96 respectively. In this manner, the Vox packet 95 can be transported across underlying UDP, IP and Ethernet layers of the network 18. This is a standard protocol encapsulation technique used by packet networks.

[0135] Referring to Figure 4C, a diagram illustrating a media multiplex Vox packet 95 encapsulated in UDP, IP, and Ethernet 97 is illustrated. In this example, the Vox

packet 95 includes a Media type field, a Media sub-type field, a length field, a message ID field, a time stamp field, a sequence ID field, and a Media payload field.

[0136] Referring to Figure 4D, the format of an indexed media payload 98 is illustrated. The indexed media payload includes a sub-type field, a length field, a message identifier (ID) field, a time-stamp field, a sequence identifier (ID) field, and field for the Media payload.

[0137] The encapsulation of Vox packets 95 into the transport packets of the underlying network allows the Media, Messages and Conversations to each be defined by a number of attributes.

[0138] When Media is created or otherwise originated on a Device 13, it is typically time-based, meaning it changes in some meaningful way over time. As a person engages in a Conversation for example, their spoken words are strung together into sentences or statements, which vary over time, and the resulting data (streams and packets) will maintain the same variance over time. Similarly video (as opposed to a still photo) as well as GPS or other sensor data will vary over time. Regardless of the type or how it is originated, the Media is segmented and placed into the payloads of a plurality of Vox packets 95. The packets are then continually, stored, transmitted, received, stored and rendered in streams (i.e., streaming media) at the transmitting and receiving Devices 13 respectively. Since each packet 95 is indexed, time-stamped, and given a sequence identifier, the individual packets can be segmented into Messages. By sequentially threading the individual Messages together, Conversations may be constructed.

[0139] One further unique aspect of the system 10 is that the media payloads generated by a Client 12 are stored in multiple locations. Not only are the payloads stored in the PIMB 30 of the generating Device 13, but also in the PIMB 85 of the Server(s) 16 and the PIMB 30 of the receiving Devices 13. This basic feature enables or makes possible much of the Voxing functionality described above and provides the system 10 with both resilience and operability, even when network conditions are poor or when a Participant of a Conversation is not connected to the network.

G. Interoperability With Underlying Telecommunication Protocols

[0140] The system 10 is intended to run or be layered over a variety of existing communication networks 18, such as the Internet, fixed PSTN type circuit networks, and mobile or cellular phone networks, or a combination thereof. The system 10 is

designed around the concept of moving many small units of information (i.e., the Vox packets) between different Clients 12 and Servers 16 in the system 10. While Vox packets may vary in size, depending on their function and payload, they all appear to be the same kind of data to the underlying network layer. In one embodiment, the system 10 has been designed and optimized for IPv4 networks such as the Internet, but other types of networks may be supported as well. For the purposes of this document, the term “IP” should be taken to mean IPv4, IPv6 or any other current or future implementation of the Internet Protocol.

[0141] Referring to Figure 5, a diagram of a Client 12 running on Device 13 and communicating with a Server 16 over a shared IP network 100 is shown. As illustrated, the Client 12 is coupled to the shared IP network 100 through a first Internet service provider A and the Server 16 is coupled to the shared IP network 100 by a second Internet service provider B. During communication, the Vox packets 95 (designed “VP” in the figure) are encapsulated within UDP/IP packets and then interleaved among other IP protocol packets as is well known in the art and transmitted across the shared IP network 100 from the Client 12 to Server 16, or vice versa. As is well known, each lower packet layer encapsulates the entire packet of the layer immediately above it. Packets can also be sent in a similar manner between two Servers 16. Messages are sent from one Client 12 enabled Device 13 to another over a shared IP network 100. At each hop, the Vox packets 95 are embedded in the underlying IP protocol and transmitted, until they reach the target destination.

[0142] The diagram of Figure 5 is merely exemplary, showing only a single Client 12 and Server 16 connected to the network 100 for the sake of illustration. In actual embodiments of the system 10, a large number of Clients 12 and one or more Servers 16 are typically connected to the shared IP network 100. It is also useful to note that the Client 12 and Server 16 do not have exclusive use of the IP network 100. In the example shown, an HTTP client, which is coupled to the network 100 through Internet provider A, can send packets back and forth with an HTTP server, coupled to the network 100 through a third Internet provider C. The system 10 does not control the manner in which the VPs embedded in the IP packets traverse the network 100. Rather, all packets that traverse and share the network 100 do so in accordance with the standard procedures of the underlying shared IP network 100.

[0143] Referring to Figure 6, a “circuit” based network 104 such as a GSM mobile phone network is illustrated. The circuit network 104 is coupled between Client 12 running on Device 13 and Server 16. Once a circuit is established between the Client 12 and Server 16, the system 10 layers Vox packets (VP1, VP2, VP3, VP4, VP5, etc.) onto the underlying packets used by the network 104 and transmits them across the network 104, creating “virtual Vox” circuit. The Vox Packets sequentially traverse the circuit network 104, typically with spacing or framing data as is well known in the art for transmitting data over a circuit network. In addition, packet construction parameters, such as the payload size and the number of header fields included may be used to exploit the lack of per-packet overhead and to increase speed and/or efficiency of data transfer across the network 104. It should be noted again that for the sake of simplicity, only a single Client 12 and Server 16 are shown connected to the network 104. It should be understood, however, that additional circuits between Clients 12 and Servers 16 as well as other components may be established concurrently through the network 104. The network 104 is therefore not dedicated for the transmission of Vox Packets, but rather may be shared with other types of network traffic.

[0144] Referring to Figure 7, a diagram illustrating communication between a first Client 12A enabled Device 13A associated with a first network A and a second Client 12B enabled Device 13B associated with a second network B is illustrated. The networks A and B further each include gateway Servers 16A and 16B respectively. The Gateway Server pair 16A and 16B facilitate communication between the two networks A and B, allowing the Devices 13A and 13B to communicate with each other. In various embodiments, the networks A and B could each be any type of network. For example, each network A and/or B could be an IP network, a circuit type network, or a wireless or cellular network (i.e., CDMA, GSM, TDMA, etc.). The Servers 16 that straddle the two networks A and B are considered gateway servers because they route traffic or serve as a “gate” between the two networks.

[0145] With the system 10, there are a several basic network interaction considerations to optimize system performance. These considerations include factors such as resolving the underlying address to which the Vox packets 95 are to be sent, the integrity of any sent Vox packets, and the management of the Maximum Transmission Unit (MTU) of a single Message that may be sent across a given network or combination of networks.

- [0146] The address of a target Client 12 needs to be known so that the underlying network delivers the Vox packet 95 to the correct location. With IPv4 networks, the address is typically an IPv4 Address, which is a 32-bit number that uniquely identifies a host within the network. For other networking technologies, the address could be some other type of identifier. IP networks use the Domain Name System (DNS) to resolve human-readable names into IP addresses, and the Address Resolution Protocol (ARP) to resolve IP addresses into physical addresses. Regardless of the underlying networking technology, the system 10 uses one of the above-mentioned or other known addressing schemes for delivery of Vox packets 95 to the correct location.
- 10 [0147] As with almost any packet-based communication system, transmitted Vox packets might not be delivered to their addressed location if the underlying network is unable to deliver the packets in which the Vox packets are encapsulated. Most packet networks do not inform transmitters when packets are dropped. Instead they rely on the transmitters and receivers to notice and compensate for any dropped packets. The system 10 is designed to use these receiver receipt report messages to coordinate this packet loss management. If the underlying network is able to inform the sender of lost or dropped packets, the system 10 utilizes this information in its retransmission protocol.
- 15 [0148] The management of MTU is the determination of the Maximum Transmission Unit (i.e., the maximum size of a single message) that may be sent across a network. For packet-based networks, the underlying network imposes the MTU. For circuit-switched networks, the MTU may be a tunable parameter for network efficiency and performance. Thus in most cases, the underlying network imposes or determines the maximum size of the Vox packet 95 that may be transmitted efficiently. For example with IP networks, packets may be fragmented if the payload exceeds the MTU, but at a substantial performance penalty. With IP over Ethernet networks, the transmitting device has an MTU of 1518 bytes, as enforced by Ethernet. The largest IP packet must leave room for the Ethernet headers. The largest UDP packet must leave room for both IP and Ethernet headers and the largest Vox protocol that may be generated on Ethernet for example is the Ethernet MTU (1518) – IP header (20) – UDP header (8) = 1490 bytes. Since the Vox protocol will have a header of its own, the actual Vox media payload will be less than 1490 bytes on an Ethernet network. For Gigabit
- 20
- 25
- 30

Ethernet, the MTU could be much larger, but would be determined using a similar formula.

[0149] In a purely packet-based network, there are two potential values for MTU, the local link MTU and the path MTU. Determining the local link MTU yields the maximum size for Vox packets to be efficiently sent out to the local network interface. The path MTU yields the maximum size of Vox packet that may be sent intact all the way to the remote node. If a sender is connected via Ethernet, the Vox packet might pass through various other systems with smaller MTUs en-route to the recipient. The smallest MTU on the path to the destination needs to be resolved and known by the sender. In the IP world, there is a standard procedure for discovering the smallest MTU called "Path MTU Discovery". For other kinds of networks, an equivalent procedure may be used. Again, since the system 10 is layered on top of other networks, any of the above MTU algorithms may be used.

H. Operation Flow Diagrams

H.1 Store and Stream

[0150] Referring to Figures 8A through 8F, a series of flow diagrams are provided to illustrate the operation of the store and stream module 24 and 84 on Clients 12 and Servers 16 respectively. Figure 8A shows the sequence of operation for a first Client 12₁ transmitting Messages to a second Client 12₂. Figures 8B and 8C illustrate the operation of the PIMB writer 28 and PIMB Reader 28 on the transmitting Client 12₁. Figures 8D and 8E illustrate the operation of the PIMB Writer 28 and PIMB Reader 26 on the receiving Client 12₂. Figure 10F illustrates a flow diagram of the Store and Stream module 84 on a server 16.

[0151] In Figure 8A, the User of Client 12₁ running on a Device 13₁ originates Media to be transmitted. The Media can be originated at the Device 13 in a number of different ways such the User creating Media by speaking into a microphone or creating video content on their Device 13. Media is also originated by a Device 13 by receiving sensor data, such a GPS information or a temperature reading. Regardless of how the Media is originated, the Media is encoded by the PIMB Writer 28 (box 130), which converts the Media into indexed media payloads and stores them in the PIMB 30 (box 132) on Client 12₁. The PIMB Reader 26 on the Client 12₁ reads the payloads out of the PIMB 30, creates Vox packets, and transmits the packets to the receiving Client 12₂ (box 134) over the network 18. Each Server 16 along the path between the

sending Client 12₁ and the receiving Client 12₂ stores the transmitted payloads in the PIMB 85 and transmits the Vox packets to the next hop (box 133). At the receiving Client 12₂, the net receive function of the PIMB Writer 28 converts the Vox packets into indexed media payloads (box 136) and stores the payloads into the PIMB 30 of Client 12₂ (box 138). The rendering module of the PIMB reader 26 on Client 12₂ renders the payload information read from the PIMB 30 into a medium suitable for human consumption, such as voice or video (box 140). Each of these steps are described in more detail below with respect to Figures 8B through 8E.

[0152] In Figure 8B, a sequence of the Encode Receive function performed by the PIMB Writer 28 (step 130 of Figure 8A) is provided in detail. In the initial step 130₁, the User of the Device 13 running the Client 12₁ originates the Media to be transmitted. As noted above, the Media may be derived by speaking into a microphone, using a camera, receiving sensor data, or by some other Media generating component. In the next step 130₂, the Encode Receiver 28a encodes the Media and creates the indexed media payloads (step 130₃), which are then stored in the PIMB 30 (step 132) by the Data storer 28c.

[0153] In Figure 8C, the sequence of the Transmit function performed by the PIMB Reader 26 (step 134 of Figure 8A) on the sending client 12₁ is provided in detail. In decision loop 134₁, the transmit function of the PIMB Reader 26 continuously checks to see if indexed media payloads that are to be transmitted have been written into the PIMB 30 and are available for transmission. If such payloads are available in the PIMB 30, the Data Prioritizer 26f prioritizes the payloads that should be sent first, using the MCMS Participant Attributes information, as illustrated in step 134₂. Information about the highest priority payloads are passed to the Packet Optimizer module 26g which runs the PQM (step 134₃), as described in more detail below with respect to Figures 9A-9C. The appropriate payloads are then retrieved from the PIMB 30 (step 134₄) by the Data Retriever 26h and converted into Vox packets 95 by the Packetizer 26i (step 134₅). The Vox packets 95 are then transmitted (step 134₆) by the Transmitter 26j over the network 18 to the receive Client 12₂, which sends back receipt reports reflecting the properties (loss, jitter, throughput) of the packets that have been received. These receipt reports provide the information necessary for the PQM to calculate the MABR for a given receiver. The aforementioned process is

repeated for each transmission loop as indicated by the return arrow from the transmit step to the top of the flow chart.

[0154] In the embodiment described above, the media is encoded, stored in the PIMB 30 and then transmitted over the network in a serial fashion. In an alternative embodiment, the encoded media can be stored in the PIMB 30 and transmitted over the network in parallel, meaning the two functions occur substantially at the same time.

[0155] In Figure 8D, the sequence for the Net Receive function (step 136 of Figure 8A) of the PIMB Writer 28 on the receiving Client 12₂ is illustrated. In the initial step 136₁, the network receiver 28d receives the Vox packets 95 over the network 18. The Data Storer 28f converts the packets into indexed media payloads (step 136₃), which are stored in the PIMB 30 (step 136₄). As the payloads are stored, the Data Quality Manager (DQM) 28g is run. The DQM 28g checks for missing or corrupted packets, ensures the eventual storage of an exact copy of the transmitted data, and sends receipt reports regarding the conditions of the network to the transmitter. Each of these functions of the DQM 28g are described in more detail below with regard to Figures 9D through 9F.

[0156] In Figure 8E, the sequence for the Render function of the PIMB Reader 26 (box 140 of Figure 8A) on the receive Client 12₂ is illustrated. In the initial step 140₁, the Data Prioritizer 26a prioritizes the indexed media payloads to be rendered as determined by the MCMS application 20 using the MTSD information as well as User status and presence information, including the User's Intentions and Attention status. The prioritized payloads are then read from the PIMB 30 (step 140₂) by the Data Retriever 26b. The PLC/Interpolator 26c performs packet loss compensation and interpolation (step 140₃) on the retrieved payloads, using known packet loss compensation and interpolation algorithms depending on which Codecs are used. In the next step, the Mixer 26d mixes (step 140₄) multiple Messages of a Conversation if two or more Participants have generated Media at the same time within the same Conversation (e.g., both are speaking at the same time). The Renderer 26e renders (step 140₅) the data stream from the Mixer 26d, generating sound, video, or other Media (step 140₆) for the recipient User.

[0157] In Figure 8F, the sequence for a Server 16 to receive Vox packets from the previous hop on the network 18, store, archive and transmit the Vox packets to the

next hop is illustrated. In the initial step, the Server 16 performs the Net Receive function of the PIMB writer (similar to Fig. 8D) to store the indexed media payloads of the received data in the PIMB 85 and archive 89 or the Server 16. The server 16 also performs the Transmit function of the PIMB writer (similar to Fig. 8C) to forward the received packets onto the next hop on the network 18. In this manner, a copy of the media generated by the transmit Client 12₁ is received, stored and transmitted at each hop along the path to the receive Client 12₂.

[0158] In the aforementioned embodiment, the writing of received indexed media is stored in the PIMB 91 of the Server 16 and transmitted to the next hop in a serial fashion. In an alternative embodiment, the received indexed media payloads can be stored in the PIMB 91 and transmitted to the next hop substantially at the same time. The storage of Media on the PIMB 30 of both transmitting and receiving Devices 13 allows for the progressive transmission and rendering of Media. On the transmit side, the Media originating on the transmitting device may be progressively transmitted over the network as it is being received. In various embodiments, the encoded Media (regardless of how it is originated) may be progressively transmitted before, after, or at substantially the same time it is stored in the PIMB 30. On the receive side, the incoming Media may also be progressively rendered as it is received over the network, provided the User has opted to Review the Media in the near real-time mode. In various embodiments, the incoming Media may be progressively rendered before, after or substantially at the same time as it is stored in the PIMB 30 of the receiving Device 13. If the received Media is to be Reviewed in the time-shifted mode, then the Media is retrieved from the PIMB 30 (or possibly a PIMB 85 on a Server 16 if replaced on the local PIMB 30) for later Review at a time designated by the User.

[0159] In the context of the present application, the term progressive or progressively is intended to be broadly construed and generally mean the continuous processing of a data stream based on availability of the data. For example, as Media is created or otherwise originated on a Device 13, the progressive encoding, storage, packetization and transmitting of that media is continuous, so long as the Media is available. As a person speaks, that Media is progressive or continuously encoded, stored, packetized and transmitted for the duration of the persons speech. When the person pauses or stops speaking, there is no media to progressively process. When the person resumes

speaking again, the progressive processing of the Media resumes. On the receive side, the Media is also progressively processed as the Media is being received (i.e., available). As the Media is received it is continuously stored. It will also be continually rendered as the Media is being received when in the near real-time mode or from storage when in the time-shifted mode. Although the above explanation was provided in the context of voice, it should be understood that all types of Media can be progressively processed in a similar manner. Also the progressive processing of Media does not necessarily have to be progressively processed in indexed order. Rather the Media is processed in the order in which it is received or by using other indexing schemes. If Media is received out of index order, in one embodiment, the Media is progressively processed in the order it was received and then organized into the indexed sequence in the PIMB 30. In an alternative embodiment, the received Media can be organized into its indexed sequence and then progressively rendered.

H.2 PQM Operation Flow Diagrams

[0160] The PQM 26g relies on a metric called the Maximum Available Bit Rate (MABR), which is a continually computed approximation of actual transmission capacity or bandwidth (i.e., a measure of the capability of the network at a given point in time) between a sending and receiving node pair. As instantaneous network conditions change, the MABR is updated. Regular measurements of network throughput, packet loss, and jitter are considered in computing the MABR. In an alternative embodiment, the MABR may also be manually set or limited based on time of day, type of network, other conditions or parameters.

[0161] The PQM also considers the Intention of the recipient(s) to optimize transmission for time-sensitivity. A transmission is considered time-sensitive if either (i) the Intention of the recipient(s) is to Review the transmission “live” or in the near real-time mode, or (ii) the recipient would like to immediately Review a Message that for some reason is not currently stored on their Device 13 (e.g., the Message was previously stored in the archive 89). The Intention of the recipient can be either inferred by the behavior of the recipient or the recipient may set or otherwise designate their Intention. On the other hand, a transmission is considered to be not time-sensitive if the Intention of the recipient is to Review the Message in the time-shifted mode. The Intention of the recipient to Review the transmission in either the live (i.e., real-time mode) or time-shifted mode at least partially defines the

“timeliness requirements” of the transmission. In various other embodiments, factors such as the urgency or priority of transmissions may also be considered in defining the timeliness requirement of the transmission.

[0162] The nodes in the network path between a sender and a receiving pair also need to be consistent regarding the status of intentions of the recipients. If one target recipient indicates timeliness, meaning they wish to Review the transmission immediately or live, then all the intermediate nodes on the network along the sender-receiver path need to have the same timeliness requirement, regardless of the requirements of other recipients. The timeliness requirement of each of the intermediate nodes is therefore dependent on the receiving nodes the transmission is being sent to. This dependency is sometimes referred to as a “union of requirements” for target nodes in the network transmission path.

[0163] The PQM further considers an Ideal Bit Rate or “IBR” for each scheduled Message payload transmission. For time-sensitive transmissions, the IBR is computed based on the packetization rate needed for substantially real time or live communication (referred to herein as the Real Time Bit Rate or RTBR). With voice for example, a packetization rate of a packet every 20 milliseconds containing 20 milliseconds of audio data is considered an acceptable IBR for conducting live conversations. The RTBR for such a system in kilobits per second would be the size of 1 second of audio payload data plus the size of all network headers that would be generated for the transmission. For video media or a combination of voice and video, the RTBR will likely be substantially higher than simply voice. For other types of media such as sensor or GPS positioning data, the RTBR will likely be lower than that of voice. For non time-sensitive transmissions, the IBR is set to a Maximum Efficiency Bit Rate (MEBR) to optimize the usage or efficiency of transmissions over the network. The MEBR is calculated by adjusting the packetization rate to its largest possible value for the underlying network. If multiple messages or payloads are to be sent between a sending and receiving pair, then an Aggregate IBR (AIBR) is considered for the transmission.

[0164] The PQM operates by sending data in a series of transmission loops for each sending and receiving pair. The transmission loops for each sending and receiving pair are independent. Any transmission on the network may affect the MABR of other

sending-receiving pairs. Accordingly, the MABR is preferably continually computed for all recipients.

[0165] Referring to Figures 9A through 9C, flow charts illustrating the sequence of operation of the PQM for a single sending and receiving pair is illustrated. In Figure 9A, the steps in determining the MABR between the single sending and receiving pair are illustrated. In Figure 9B, a flow chart illustrating the steps for calculating the AIBR for each transmission loop for the single sending and receiving pair are illustrated. In Figure 9C, a sequence for determining the amount of data to transmit between the sending and receiving pair per loop is illustrated. The processes illustrated in the three diagrams run simultaneously and interact with one another, as described in more detail below.

[0166] In Figure 9A, a flow chart 50 for calculating the MABR for the network interface between the sending and receiving pair is shown. In the initial step 50₁, the network interface between the sending and receiving pair is monitored. The sender periodically receives reports, which contain information regarding the status of the network connection at the receiver in step 50₂. The reports include information regarding the current status of data throughput 50₃, packet loss 50₄, and jitter 50₅ as observed by the receiver at the network interface. In step 50₆, the MABR is calculated at the sender based on these observations contained in the reports. By monitoring or observing the data in these reports, the MABR value is continually adjusted based on current network capabilities or conditions between the sending and receiving pair. As network capabilities become more favorable for data transmission, the MABR will increase. If the network capabilities become less favorable for transmission, the MABR will decrease, potentially all the way to zero for an unusable network. The reports are similar to the packet loss reports generated by nodes in TCP networks, but additionally include throughput and jitter information as well.

[0167] If there are several network interfaces between the sending and receiving pair, an MABR is calculated for each interface for which a receipt report is received. If no traffic has been recently sent on the network, or no receipt reports have been received, the MABR may not reflect current network conditions. However, since receipt reports are continually generated by receivers while data is transmitted, the sender's MABR metrics will quickly converge to a more accurate value.

[0168] Referring to Figure 9B, a flow chart 52 illustrating the steps for calculating the AIBR for a transmission loop is illustrated. In the initial step 52₁, the Messages with Media (by which we mean portions of the time indexed media that belongs to this Message) ready to be transmitted between the sending and receiving pair in the current loop are ascertained. A list of Messages with Media is then built 52₂. For each Message in the list 52₃, the time-sensitivity or timeliness requirement of each Message is considered 52₄. If a particular Message is not time-sensitive, then the IBR is set to the Maximum Efficiency Bit Rate (MEBR) 52₅. On the other hand, if a Message is time-sensitive, then the IBR is set to the Real Time Bit Rate (RTBR) 52₆. In the next step 52₇, the previously computed IBRs for each of the Messages in the list are summed together, resulting in the Aggregate Ideal Bit Rate (AIBR) 52₈ for the transmission loop. As signified by the return arrow 52₉, the above-described process is repeated for each transmission loop between the sending and receiving pair.

[0169] Referring to Figure 9C, a flow chart 54 illustrating the sequence for determining the rate of data to transmit between the sending and receiving pair per transmission loop is illustrated. In the initial step 54₁, the MABR (as calculated in Figure 9A) is compared to the AIBR (as determined in Figure 9B) for the next transmission.

[0170] If the MABR is greater than or equal to the AIBR, then all the Messages identified as ready for transmission in the loop are packetized at the IBR rate 54₂ and transmitted 54₃.

[0171] On the other hand, if the MABR is less than the AIBR, then a series of procedures are applied to so that the PQM meets its goals of the timely delivery of an adequate copy of the data, the efficient use of available bandwidth, and/or adjustments to the payload content and quality, packet size, and transmission rate to meet current network conditions. In other words, the number of bits used to represent the transmitted Media when the MABR is less than the AIBR is reduced until the amount of data loss (i.e., throughput) is reduced to an acceptable level.

[0172] In the initial step, the Messages in the list are reviewed for time sensitivity 54₄. If there are no time sensitive Messages, then the bit rate is set to the MABR 54₅, and the Messages are transmitted 54₃.

[0173] If the list includes time-sensitive Messages, the bit rate allocated for the non time-sensitive Messages is reduced 54₆ until it meets the MABR limits. If reducing

the bit rate all the way to zero is insufficient to meet the MABR, then these non time-sensitive Messages are removed from the list of Messages to be transmitted in the loop. If the bit rate has been reduced so that it is less than or equal to the MABR, then the remaining Messages are Packetized and transmitted 54₃.

5 [0174] If the removal of non time-sensitive Messages was not sufficient to meet MABR, then a procedure for sending a reduced bit version relative to the IBR version of just the time-sensitive Messages is performed for the purpose of increasing the ability of the recipient to Review the Messages upon receipt so the Conversation may continue in the near real-time mode. In one embodiment, generating the reduced bit
10 version of the time-sensitive Messages involves using fewer bits relative to the IBR when packetizing the Media of the time-sensitive Messages to be transmitted. For example, the number of bits used per payload may be progressively lowered until the available bandwidth on the network is met. This procedure may be accomplished by either adjusting the setting of Codecs, using different or lower quality Codec (or
15 Codecs), or any combination thereof. In another embodiment, generating a reduced bit version of the time-sensitive Media contained in the Messages to be transmitted may involve the use of any of a number of known compression algorithms. If the reducing the number of bits is sufficient, meaning the MABR limit is met, then the Messages are transmitted 54₃. Regardless of which embodiment is used, an attempt is made to
20 transmit the payload data as fast as possible by sending fewer bits during the transmission loop. By reducing the quality of the payload, the transmission by sending fewer bits in a given period of time, the Conversation may be maintained in the near real-time mode.

[0175] If reducing the bit rate using any of the above-mentioned techniques still does
25 not meet the MABR, then the packetization interval for transmitting the time-sensitive Messages is increased 54₈. With this procedure, the header-to-payload ratio is increased, which lowers the overall bit rate but introduces latency (i.e., a delay in the delivery of the transmission to the recipient). If this procedure results in the reduction of the AIBR to less than or equal to the MABR, then the transmission 54₃ occurs.

30 [0176] If after changing the packetization interval the MABR is still not met, then the bit rate may be progressively lowered 54₉ to be within the MABR limit. When the bit rate is lowered in this manner, time-sensitive messages are sent at a rate that is not sufficient to maintain a live conversation. Therefore, the Conversation is forced out of

“live”. If the network is down or conditions are very poor, it is possible that no data transmission may occur. Again, the aforementioned sequence is repeated for each transmission loop 54₁₀ between the sending and receiving pair. As conditions on the network improve, the Conversation may resume in the live or near real-time mode by decreasing the packetization interval and/or using more bits to represent the Media to meet network conditions.

[0177] If there are multiple network interfaces between the sending and receiving pair, the sequence described in relation to Figure 9C is performed for each interface for which receipt reports are available. In various embodiments, the sender may contain logic to distribute the transmission load among the multiple interfaces. In different examples, the payloads can be sent only on one interface, while in other embodiments, some or all of the interfaces may be used.

[0178] The aforementioned description pertains to any sending and receiving pair in the system 10. In most situations, the sending and receiving pair will include a Client 12, enabled Device 13 and Server 16, two Servers 16, a Server 16 and Client 12 enabled Device 13, or even possibly two Clients 12 respectively. If a sending node is transmitting to two (or more) receiving nodes at the same time, the above mentioned sequence as described in relation to Figures 9A-9C occurs concurrently for each receiving-sending pair.

20 **H.3 DQM Operation Flow Diagrams**

[0179] The DQM 28g determines if data received at the Client 12 is either corrupted or if there are missing packets. In addition, the DQM 28g of a receiving Client 12 generates of the receipt reports, which are sent back to the transmitting node on the network. The DQM 28g also runs a background process to ensure that an exact or complete copy of transmitted data is eventually received and stored. These functions are described below in Figures 9D through 9F respectively.

[0180] Referring to Figure 9D, a flow diagram illustrating the operation of the DQM 28g for checking for missing and/or corrupted data is illustrated. In the initial step 56₁, the DQM 28g checks for corrupted packets using well-known techniques, such as CRC or similar integrity checking mechanisms. If a packet is corrupted, that packet is treated as missing 56₂. The DQM 28g next ascertains if any packets are missing 56₃. If an out of sequence packet is not received after a predetermined period of time, it is assumed to be missing. The DQM 28g notes any missing or corrupted packets 56₄ in

the DNQS 32. On the other hand if no corrupted or missing packets are detected, the DQM 28g determines if the quality of the received data was intentionally degraded (i.e., a reduced bit rate representation of the Media) by the sender 56₅ for the purpose of saving bandwidth. The degraded quality is noted in the DNQS 32 56₆. Regardless
 5 if the quality of the received data is degraded or not, receipt information (e.g., a packet sequence number, time stamp, and the network address of the next node in the network the packet is to be sent) of the data is stored 56₇ in the DNQS 32. The aforementioned process is continually repeated, as signified by the return arrow to the start of the flow diagram.

10 **[0181]** As a result of the process detailed in Figure 9D, information regarding the receipt of non-degraded (i.e., full bit rate representations of Media) packets, the deficiency of degraded (i.e., reduced bit rate representations of Media) quality packets, and missing or corrupted packets, are all stored in the DNQS 32. As Media is received, the DNQS 32 maintains up-to-date information regarding the status of the
 15 Media.

[0182] Referring to Figure 9E, a flow diagram illustrating the operation of the receipt report generator function of the DQM 28g is illustrated. In the initial step, the DNQS 32 is periodically scanned 58₁ to determine if there is any Media for which a receipt report needs to be generated 58₂. If the answer is no, then the above scanning process
 20 is repeated. On the other hand if Media is identified, then the process determines if the Media is time-sensitive 58₃, meaning either the User intends to Review the Media live or the user would like to immediately Review Media that is not stored on their Device 13.

[0183] If not time-sensitive, then the recipient informs the sender to set the
 25 retransmission priority (as defined below) to low 58₄. If time-sensitive, then the amount of packet loss is considered 58₅. If the amount of packet loss is outside a usable quality range, the retransmission priority is set to high 58₆. As noted above, if the amount of packet loss is too large, the recipient may not be enabled to Review the Media upon receipt. If the quality is within an acceptable range, meaning the quality
 30 of the transmission (i.e., the reduced bit rate representation) is sufficient that it can be understood when rendered, then the priority for the sending of the receipt report is set to low 58₄. Regardless if the recipient is Reviewing upon receipt or not, a receipt report is sent 58₇, the DNQS 32 is updated 58₈ and the Network Quality Manager

(NQM) 28h is updated 58₉. The retransmission requests defined in step 58₄ is therefore conditional based on time-sensitivity. The transmission request defined in step 58₆ is conditional on both time-sensitivity and quality.

[0184] The retransmission priority informs the PQM 26g of the sender to properly
5 prioritize the transmission rate for the Media that requires retransmission. For example when the retransmission priority is set to high, then the sending PQM 26g of the sender should send any retransmissions before any new Media. If the priority is low, the PQM 26g should send the retransmitted Media after any new Media.

[0185] The aforementioned process is continuously repeated so that receipt reports
10 are generated as Media is received. If the sender does not receive receipt reports in a timely manner, the PQM 26g of the sender will reduce the transmission rate, eventually stopping the transmission if no receipt reports are received.

[0186] Referring to Figure 9F, a flow diagram illustrating the sequence for requesting
of missing or degraded Media is illustrated. In the initial step 60₁, the DNQS 32 is
15 periodically scanned for missing or degraded Media 60₂. If there is no missing or degraded Media, then the above defined scan is periodically repeated. In one embodiment, high priority request means an immediate request for retransmission and a low priority request means a deferred request for retransmission, typically until
20 bandwidth in excess of what is needed for high priority retransmissions and/or time-sensitive Media is available. In addition, the fulfilling of competing low priority requests at the sending node may be processed in accordance with any one of a number of different schemes. For example, requests for retransmission of low priority Media may be fulfilled on a first come-first serve basis, time-sensitive Media always before non time-sensitive Media, or vice-versa.

[0187] Media is considered missing if an out of sequence packet does not arrive after
25 a predetermined threshold period of time 60₃. If the packet arrives before the threshold, then it is no longer considered missing. On the other hand if the packet does not arrive after the threshold is exceed, then it is deemed missing. With missing packets, a low priority request for retransmission is made 60₄ and the time of the request is noted 60₅ in the DNQS 32. This process is repeated until the missing packet
30 is received. When the missing packet arrives and the corresponding Media is available in the PIMB 30, the missing status of the Media is removed from the DNQS 32. The

retransmission request defined in step 60₄ is therefore conditional based on whether the Media is determined to be missing.

[0188] If the Media degraded to a reduced bit rate representation, the DQM 32 determines if the Media is part of a live or real time Conversation 60₆. If not, a request for a full quality or bit rate copy of the degraded Media is made 60₇, the full quality Media is designated as missing 60₈ and the request time is noted 60₉, in the DNQS 32. If the Media is part of a live Conversation, then no action is immediately taken in order to preserve network bandwidth. When the Conversation transitions out of the live mode, then the steps 60₇ through 60₉ are performed to ensure that a full quality (i.e. an exact or complete) copy of the degraded Media is eventually received. When the data becomes available in the PIMB 30 of the recipient Client 12, the degraded status of the associated Media is removed from the DQNS 32. The transmission request defined in step 60₇ is therefore conditional on whether the Media is both degraded and not part of a live conversation.

[0189] The aforementioned process is continually repeated, as signified by the return arrows from 60₅ and 60₉ to the top of the flow diagram at 60₁. By repeating the process outlined in Figure 9F, exact copies of all transmitted Media is eventually stored in the PIMB 30 of the receiving Device 13. In this manner, the storage of exact copies of transmitted Media is guaranteed at the recipient Device 13.

I. Real Time Synchronization for Voice Communication

[0190] Referring to Figure 10, a diagram illustrating the real-time synchronization of the Media of a Conversation between a first Participant using Device 13A and a second Participant using Device 13B is illustrated. The progressive nature of the store and stream module 24 ensures that the Media of the Conversation is progressively stored at both Device 13A or Device 13B as the Media is created. On the transmit side, the operation of the PQM 26g (Figures 9A through 9C) progressively transmits and stores in the local PIMB 30 any Media created using the sending Device 13 as the Media is created. On the receiving Device 13, the DQM 28g progressively stores in the local PIMB 30 any received Media as it is being received and rendered. As a result, both Participants of the Conversation have synchronized copies of the Media of the Conversation as the Media is being created.

[0191] For the sake of clarity, Figure 10 shows only two Devices 13A and 13B engaged in the Conversation. It should be understood that the real-time

synchronization of the Media may be realized regardless of the number of Participants in the Conversation. In addition, the network 14 in the figure is shown as a generic networking cloud for the sake of simplicity. Again, it should be understood that the network 14 may include one or more Servers 16, where the store and stream module 84 and PIMB 85 of each intermediate hop between the two Devices 13A and 13B stores a copy of the Media of the Conversation and optionally performs real-time synchronization with Devices 13A and 13B.

[0192] Referring to Figure 11, a flow chart illustrating the sequence for the near real-time synchronization of Media of a Conversation between Device 13A and Device 13B is illustrated. In this example, it is presumed that the Users of Device A and Device B are engaged in the Conversation in the near real-time mode. The Devices 13A and 13B may be engaged in the Conversation either directly or through one or more intermediate Servers 16 (not illustrated).

[0193] After the Conversation is initiated (1102), the Users of Device 13A and Device 13B will each typically create Media. On the left side of the figure, the sequence for synchronizing the PIMB 30 of Device 13A is illustrated. On the right side, the sequence for synchronizing the PIMB 30 of Device 13B is illustrated. The sequence is essentially the same on both sides of the diagram.

[0194] During the course of the Conversation, the User of Device 13A will generate voice Media (1104A). As this Media is created, it is encoded and stored in the indexed format in the local PIMB 30 of Device 13A, as described above with regard to Figure 8B (1106A).

[0195] Device 13B notifies Device 13A of the status of all the locally created voice Media originating at Device 13B pertaining to the Conversation (1108A). In various embodiments, the notification of new Media may take a variety of different forms. For example, the notice may occur: (i) when Device 13B starts generating and transmitting Media associated with the Conversation for the first time (ii); when Device 13B logs in; (iii) each time Media pertaining to the Conversation is created at Device 13B and transmitted to Device 13A; (iv) when Device 13B periodically notifies Device 13A of the state of all the Media created using Device 13B pertaining to the Conversation; (v) Device 13A periodically asks Device 13B for the state of the Media created and stored at Device 13B, or (vi) Device 13B notifies Device 13A with a synchronization update message that Media is available either just prior to or

concurrent with the transmission of the Media. In yet another embodiments, (vii) the Media of the Conversation may be stored at one or more Servers 16 hops between the Device 13A and Device 13B. The various notices described above with regard to (i) through (vi) may occur between the Servers 16 and the Device 13A and/or 13B rather than directly between the Devices 13A and 13B. Lastly, the notice may involve some combination of (i) through (vii). As the Media is transmitted (as described above with respect to Figure 8C), it is sequenced and time-stamped, allowing Device 13A to determine which Media it has already received and which media it may be missing.

[0196] In response to the notification, Device 13A requests all the voice media associated with the Conversation not already stored in the PIMB 30 of Device 13A (1110A). This is accomplished by following the routine described above with regard to Figure 8D, including running the DQM 28g as described in relation to Figures 9D-9F. Depending on circumstances, the requested Media could be Media with a specific time-stamp or within a time-stamp range. For example, if Device A is in receipt of a less than complete copy of previously sent Media, then the request identifies only the Media needed to backfill and complete the copy of the Media at Device A. Alternatively, during the course of an ongoing Conversation, the request may not have a specific end-point. Instead, the request from Device A to Device B is open-ended and Device B sends the new Media as it is being created.

[0197] Device 13B in response sends the requested Media along with any new Media associated with the Conversation as it is being created (1112A), using the sequence described above with respect to Figure 8C. In response, Device 13A locally stores all the received Media in its PIMB. As a result, Device 13A is able to maintain a complete copy of the Media of the Conversation as the Media is created on Device 13B.

[0198] When Device 13B creates voice Media, the sequence 1104B through 1114B is executed. The sequence 1104B through 1114B is the complement of 1104A through 1114A, and therefore, is not described in detail herein. The net result of the two sequences (1104A through 1114A) and (1104B through 1114B) is that both Device 13A and Device 13B each maintain synchronized copies of the Media of the Conversation as the Conversation occurs in real-time.

[0199] The synchronization of the Media of the Conversation at Device 13A and Device 13B occurs even when the MABR is less than the AIBR. As noted above with respect to Figure 9C, a number of procedures are implemented to maintain the “live-

ness” of the Conversation when the MABR is less than the AIBR, including transmitting only time-sensitive media, using fewer bits to represent the Media per unit of time, increasing the packetization interval, and reducing the rate the Media is transmitted to meet conditions on the network 14. These procedures have the advantage of extending the ability to maintain the live-ness of the Conversation, but at the expense of not transmitting an exact or full copy of the Media as it was originally encoded. As a result, synchronization is maintained, but with a less than exact copy of the originally encoded Media at the receiving Device. With the DQM (Figures 9D through 9F) at the receiving Device, replacements for the missing, corrupted and less than full bit rate versions of the Media are eventually transmitted and stored at the receiving Device. This backfill procedure eventually results in the complete synchronization of the Media of the Conversation on Devices 13A and 13B. Thus, under certain conditions, real-time synchronization of a complete or exact copy of transmitted Media may not be possible. Rather the synchronization may be a work in progress where a complete copy of the Media is received only after the backfill process is completed.

[0200] Although the above explanation has been described in the context of a voice Conversation, it should be understood that Media of the Conversation could involve any type of Media in addition to voice, including video, GPS information, or other sensor data. In addition, the Conversation does not necessarily have to be limited to two Users or Participants. The near real-time synchronization of the Media of a Conversation occurs regardless of the number of Devices 13 participating in the Conversation.

J. Distributed Services Architecture

[0201] In the above discussion, the network infrastructure 14 has been represented as a single Server 16 serving multiple Client 12 enabled Devices 13. In an alternative embodiment, the networking infrastructure is a distributed network of Servers 16.

[0202] Referring to Figure 12, an example of a distributed services network 14 according to the present invention is illustrated. The distributed network 14 includes a plurality of Servers 16. Each of the Servers 16 may be configured to implement one or more distributed network related functions. A given Server 16 may be configured as a home Server, a gateway client Server, a gateway Server, a directory Server, an access Server, or any combination thereof. Regardless of how many functions a particular

Server 16 is configured to perform, each shares the architecture as illustrated in Figure 3 and includes a Store and Stream module 84 and PIMB 85, with one possible exception. If a Server 16 is configured solely as a directory Server, then the module Store and Stream 84 and PIMB 85 are not required.

5 [0203] Each User is assigned a specific Server 16 as their home Server. A home Server is where the Media of each of the User's Conversations is stored on the network 14. For each Conversation, the stored Media includes the Media contribution of all the Participants, including the contributions of the User.

[0204] In one embodiment, home Devices 13 rely on the well known Domain Name
10 System (DNS) for locating their home Servers 16 and for locating the home Servers 16 of other Users. In alternative embodiments, the Lightweight Directory Access Protocol (LDAP) is used. As should be obvious to those skilled in the art, any type of networked directory lookup system may be used for locating home Servers 16.

[0205] Home Servers 16 also manage and perform home Server migration. Home
15 Servers 16 are usually located on the Network 14 topology in a convenient place for where a User ordinarily accesses the network 14. If a User accesses the network 14 from a different part of the network topology for some period of time for whatever reason, the home Server 16 for that User may be moved to another Server 16 that is more optimal in the network topology for that User. When this occurs, the User's
20 persistently stored Media is migrated from the previous to the new home Server 16 over the network 14. This migration process may be initiated automatically after a predetermined period of time after the User has moved, or it can be implemented manually.

[0206] Gateway Servers 16 are gateways between the network 14 and other
25 communication networks, as illustrated in Figure 7 for example. Gateway Servers 16 straddle the two networks and are responsible for performing any mapping, trans-coding re-encapsulation, security, or other translation functions as necessary when routing traffic between the two network domains. When the two networks have different packetization requirements for example, the gateway Server 16 re-packetizes
30 the packets from the format used by one network into the format used by the second network and vice-versa.

[0207] Gateway client Servers 16 provide intermediation between the distributed services architecture 14 and non-Client 12 enabled communication devices 1202.

Gateway client Servers 16 perform on the network 14, on behalf of the devices 1202, the functionality of (i) the MCMS application 20 (including MCMS, and optionally MCMS-S and/or MCMS-C) and (ii) the Store and Stream and persistent storage operations of module 84 and PIMB 85. Specifically, Gateway client Servers perform
5 the “PIMB Reader and Render” function (Figure 8E), “PIMB Writer Net Receive” function (Figure 8D), “PIMB Writer Encode Receive” function (Figure 8B) and “PIMB Reader Transmit” function (Figure 8C) so that the Media of Conversations is stored in the PIMB 85 of the Server 16 on behalf of the non-Client 12 enabled device 1202. Examples of non-Client 12 enabled devices may include standard landline or
10 PSTN telephones, mobile or satellite phones, tactical radios, or any other known legacy communication device. In addition, non-Client 12 enabled communication devices 1202 may include non-legacy devices specifically made to operate in cooperation with Gateway client Servers 16. For more details on gateway client Servers, see U.S. application Serial Numbers 12/206,537 and 12/206,548, each filed
15 on September 8, 2008 and each entitled “Telecommunication and Multimedia Management Method and Apparatus”, both of which are incorporated herein for all purposes.

[0208] The access point where a User connects to the network 14 through an underlying network 18 is called an access Server 16. The access Server 16 for a User
20 may change. For example, a User may disconnect from the network 14, move to a new physical location, and then reconnect through a different Server 16. Also the access Server 16 of a User may migrate from one access Server 16 to another. For example if a User has wireless access to the network 14 through a first node of an underlying network 18 while traveling in a car, the access Server 16 may change as
25 the User moves out of the range of the first node of the wireless network 18 into the range of a second node of the wireless network 18. If the access point to network 14 for the two nodes of the underlying wireless network 18 is different, than the access Server 16 for that User will change. The ability of a User to access the network 14 through an arbitrary access Server 16 improves network efficiency by optimizing the
30 communication between the User and other nodes on the network 14.

[0209] Directory Server 16 provides routing information for the rest of the network 12. When a User joins the network 14, the access Server (i) notifies the directory Server 16 that the User has joined and (ii) identifies itself as the access Server 16 for

that User. The directory Server 16 aggregates (i) and (ii) for all Users connected to the network 14, along with the (iii) home Server 1202 of each User into a directory. The directory Server 16 is continually updated. As Users join or leave the network, or as the access Server 16 and/or the home Server 16 for a User changes, the directory is
5 updated. In this manner the directory always provides current or up-to-date information for routing Messages and other network traffic between Users across the network 14 or users located on other networks through gateway Servers 16.

[0210] If a User's home Server 16 and access Server 16 are the same, then that Server 16 persistently stores all Media transmitted or received by that User. For Media
10 created by the User for transmission and for Media sent to the User, the Server 16 performs the functions described above with relation to Figure 8F, including the PQM (Figures 9A-9C) and PQM (Figures (9D-9F) operations. In this manner, copies of the Media transmitted or received by the User are persistently stored at the home Server 16 of the User. In addition, any less than complete or full copies, missing or corrupted
15 copies of Media are eventually received through the backfill operation of the PQM 26g and DQM 28g.

[0211] If an access Server 16 is not the home Server 16 for a User, the transmitted and received Media is at least temporarily stored at the access Server 16 using the same operation described in relation to Figure 8F, including the PQM (Figures 9A-
20 9C) and PQM (Figures (9D-9F) operations. In one embodiment, the Media is stored at the access Server 16 only until the confirmation of the delivery and persistent storage of the Media to the User's home Server 16 and Device 13. In an alternative embodiment, the Media is persistently stored at the access Server 16 and the home Server 16, even after the delivery of the Media is confirmed. Different schemes for
25 deleting the Media at the access Server 16 after the confirmation of the delivery to the home Server 16 and/or Device 13 may be employed, depending on the application and network topology.

[0212] A home Server 16 is also responsible for archiving the Media of the Conversation in its archive 89. Media may be archived for a number of reasons, such
30 as restoring the Media to a Device 13 after it was locally deleted, or the sending of the Media to a new Device 13 after authentication by a User on that Device 13.

[0213] In the embodiment illustrated in Figure 12, a number of Client 12 enabled Devices 13 are connected to the network 14 through various access, home, and

gateway Servers 16. In addition, non-Client 12 enabled devices 1202 are also connected to the network 14 through a gateway client Server 16. As noted above, each Server 16 may implement one or more of the above-listed functions. For example, in situations where a Device 13 connects to the network 14 through its home Server 16, then that Server performs the dual role of being an access Server and home Server for that Device. In situations where a Device 13 accesses the network 14 through a Server 16 other than its home Server 16, then the access and home Servers 16 for that device are different. In yet other situations, home Servers and/or access Servers 16 may also function as gateway Servers, gateway client Servers, and/or directory Servers as well.

When a Server 16 acts as a hop on the network 14 for transmitting Messages, regardless of which function(s) the Server 16 is configured to implement, the Server 16 stores and transmits the Media as described above with regard to Figure 8F, including performing the PQM (Figures 9A-9C) and PQM (Figures (9D-9F) operations (except for possibly a Server 16 configured solely as a directory Server).

Each of the Servers 16, regardless of the function or functions performed, are also multi-tenant, meaning they can each support or serve multiple Users at the same time.

[0214] It should be noted that the specific arrangement of Servers 16 and the function or functions as described or shown in Figure 12 is only exemplary and is provided for the sake of illustration. In actual embodiments, the number of Servers 16, the one or more functions performed by each, and the actual configuration of the distributed network 14 may vary in nearly an infinite number of variations.

[0215] In an alternative embodiment, the network 14 may be configured without a dedicated directory Server 16. In such embodiments, access Servers 16 are configured to assume some or all of the directory and routing functionality. Some or all of the access Servers 16 in the network 14 may be configured to maintain routing tables, including “default” routes, for locating and sending Messages to other access and home Servers 16 and Users.

[0216] In another embodiment, a Client 12 enabled Device 13 may act as a Server 16 for other Devices 13 or 1202. For example, a vehicle-based Client 12 enabled radio Device 13 could be a client for the driver of the vehicle as well as a Server for all nearby radios. In another example, a computer could act as not only a Client 12 for a local User, but also as a Server for other Client 12 enabled computers connected on the local network.

[0217] In yet other embodiments, access Servers 16 may provide proxy access to home Servers 16 and/or directory Servers 16 when the network connection to the home Server and/or directory Server is degraded or unavailable. Access Servers also provide current optimal routing information by locating the appropriate home Servers, other access Servers and Client enabled Devices 13 and maps the most efficient path between Client enabled Devices 12 and these Servers 16.

K. Network Routing

[0218] As Conversations occur between the various Client 12 enable Devices 13 and/or non-Client 12 enabled devices 1202, individual Messages are typically transmitted from one or more Devices 13 and/or 1202 to one or mores Servers 16, from Server(s) 16 to Server(s) 16 across the distributed network 14, and then from Server(s) 16 to one or more Devices 13 and/or 1202. The Messages traverse the network 14 in the form of Vox packets 95, embedded in whatever type of packets are used by the underlying network or networks. In various embodiments, the underlying network may rely on the Internet Protocol (IP), a Real Time Protocol (RTP), or any other type of protocol, for the transmission of network traffic. In routing the individual Messages across the distributed architecture 14, a procedure that reduces traffic and latency, while maintaining simplicity, is preferred.

[0219] Referring to Figure 13, a simple distributed network 14 for illustrating a routing procedure used in one embodiment of the present invention is illustrated. In this embodiment, the network 14 includes access Servers 16 labeled "Access Server 1" and "Access Server 2", Client 12 enabled Devices 13₍₁₎ through 13₍₅₎ a number of home Servers 16, and a directory Server 16. Devices 13₍₁₎ through 13₍₄₎ are connected to the network 14 through Access Server 1. Device 13₍₅₎ is connected to the network 14 through Access Server 2. Home Servers 1, 2 and 5 are the home Server 16 for Devices 13₍₁₎, 13₍₂₎, and 13₍₅₎ respectively. Devices 13₍₃₎ and 13₍₄₎ share a home Server 16. A sixth home Server 16 is also provided for a sixth Devices 13₍₆₎, which is purposely not illustrated because it is currently disconnected from the network 14 in this example.

[0220] When Device 13₍₁₎ initiates a Conversation with Devices 13₍₂₎ through 13₍₆₎, it sends a request to Access Server 1, which notes that it already has connectivity with Devices 13₍₂₎ through 13₍₄₎. Access Server 1 also requests from the directory Server 16

the status of Devices 13₍₅₎ and 13₍₆₎ and the Home servers for Devices 13₍₁₎ through 13₍₆₎. The directory Server 16 responds by notifying Access Server 1 that Device 13₍₅₎ is online through Access Server 2, and Device 13₍₆₎ is not online. The directory Server 16 also notifies Access Server 1 that the home Servers for participants 13₍₁₎, 13₍₂₎,
5 13₍₃₎ - 13₍₄₎, 13₍₅₎ and 13₍₆₎ are Home servers 1, 2, 3 - 4, 5 and 6 respectively.

[0221] When Device 13₍₁₎ sends Messages related to the Conversation, the Messages are sent to the Access Server 1. In turn, the Messages are directly forwarded to Devices 13₍₂₎ through 13₍₄₎ since these Devices are connected to Access Server 1. In addition, Access Server 1 notes that the remaining transmission destinations,
10 specifically Home servers 5 and 6 and Devices 13₍₅₎, have the same next hop, which is Access Server 2. Since these destinations all share the same next hop, Access Server 1 sends only a single copy of each Message to Access Server 2. The Messages are then sent by access Server 2 to Device 13₍₅₎, home Servers 5 and home Server 6. Each of the home Servers 1 through 6 retain the messages for the Devices 13₍₁₎ through 13₍₆₎
15 respectively. When Device 13₍₆₎ comes online, home Server 6 forwards the Messages to Device 13₍₆₎, enabling the Device to receive the Media of the Conversation. If Devices 13₍₅₎ and 13₍₆₎ had different next hops on the network 14, then only a single copy of the Messages would be sent to each next hop. Messages generated by any of the other Devices 13₍₂₎ through 13₍₆₎ are routed across the network 14 in a similar
20 manner as described above.

[0222] The above-described routing procedure provides a number of advantages. By routing Messages directly between Devices 13 identified with the same access Server 16 through that same access Server, network traffic and latency is reduced. Also by forwarding only a single copy of Messages, as opposed to multiple copies to multiple
25 recipients sharing the same next hop on the network, network traffic and latency is again reduced. By consolidating Message transmissions to multiple recipients along the same hop to a single transmission and duplicating or “fanning out” Messages only when there are multiple next hops, overall network traffic and latency is significantly reduced.

[0223] In an alternative embodiment, access Servers 16 may be configured to allow
30 Devices 13 connected to the network 14 through the same access Server 16 to communicate directly with one another. With this embodiment, any Messages directly sent between the two (or more) Devices 13 still have to be sent to the common access

Server 16, which in turn, forwards the Messages to the home Servers for the same Devices 13. This arrangement provides an improvement in latency for communication between Devices 13 connected to the same access Server 16, but at the expense of increased complexity on both the Devices 13 and the access Server 16. Since all
5 Messages still have to be sent to the access Server 16, network traffic is essentially doubled.

L. Real-time Synchronization Across the Distributed Services Architecture

[0224] One advantage of using Servers 16 is the ability to maintain near real-time
10 synchronization of the Media transmitted across the distributed network 14. Referring to Figure 14, an example of a network 14, which is configured on top of one or more networks 18 (not shown), is illustrated. In this example, two Client enabled Devices 13₍₁₎, 13₍₂₎ and a non-Client 12 enabled device 1202 are engaged in a voice (and optionally other Media types) Conversation across the network 14. The network 14
15 includes access Servers 16₍₁₎ and 16₍₂₎, which provide network access for Devices 13₍₁₎ and 13₍₂₎ respectively. Non-Client 12 enabled device 1202 connects to the network 14 through the Server 16 configured as a gateway client. In addition, a directory Server 16 is provided. Each of the Servers 16 includes a Store and Stream module 84 and PIMB 85, except perhaps the directory Server 16, which may
20 optionally include module 84 and PIMB 85.

[0225] As Messages of the Conversation are transmitted between each sending and receiving pair on the network, the Store and Stream module 24 (for Devices 13₍₁₎ and 13₍₂₎) and 84 (for each of the Servers 16) run the sequence described above with regard to Figure 11. As a result, the Media of the Conversation is synchronously
25 stored in the PIMB 85 of each Server 16 hop on the network 14 between the legacy device 1202 and Devices 13₍₁₎ and 13₍₂₎. The Media is also synchronously stored in the PIMB 30 of the Devices 13₍₁₎ and 13₍₂₎.

[0226] In one embodiment, the Media is synchronously and persistently stored at each Server 16 hop on the network 14. In an alternative embodiment, the Media of the
30 Conversation is persistently stored on the network 14 only on the gateway client Server 16 for the legacy device 1202 and the home Servers 16₍₁₎ and 16₍₂₎ for Devices 13₍₁₎ and 13₍₂₎ respectively, but not the access Servers 16₍₁₎ and 16₍₂₎. In the latter embodiment, the Media is no longer stored on the access Servers 16₍₁₎ and 16₍₂₎ once

it has been confirmed that the Media is persistently stored on home Servers 16₍₁₎ and 16₍₂₎ and/or Devices 13₍₁₎ and 13₍₂₎ respectively.

[0227] As the Media migrates through the network 14 from its source to destination, the Media is synchronized between each sender-receiving pair. Each node on the network, including the Client enabled Devices 13₍₁₎, 13₍₂₎ and each intermediary Server 16, is synchronized “hop-by-hop”. By performing the synchronization of the Media between each sending-receiving hop on the network 14, as opposed to just the sending and receiving communication Devices 13₍₁₎ and 13₍₂₎, the resiliency and redundancy of the Media across the network 14 is increased.

[0228] Although synchronization across the network 14 has been described in the context of a voice Conversation, it should be understood that Media of the Conversation could involve any type of Media, including voice, video, GPS information, or other sensor data. In addition, the Conversation does not necessarily have to be limited to three participants as described above. The near real-time synchronization of the Media of a Conversation occurs regardless of the number participants. Nor does the Media have to be transmitted in the context of a Conversation. Media not pertaining to a Conversation can be synchronized across the network 14. Lastly, it again should be understood that the particular network configuration illustrated in Figure 14 is merely illustrative for explaining the real-time synchronization across the network. The particular configuration shown should not be construed as limiting in any way.

[0229] It should be noted that in the above discussion with regard to Figures 8A through 8F, Figures 9A through 9F, and Figures 10 through 14, the Media has been described as indexed in a time-based format. It should be understood that time-indexing the Media is exemplary and should not be construed as limiting the present invention. On the contrary, any indexing format may be used, such as indexing by bit or byte sequence, hashing type indexing, binary-tree type indexing, file-directory based indexing, or any combination thereof. The term “indexing” therefore should be broadly construed to cover any indexing scheme.

M. Client and Server Hardware

[0230] Referring to Figure 15A, a block diagram 140 illustrating the hardware of a Device 13 used for storing and executing the Client application 12 is shown. The hardware includes a CPU 142, main memory 144 and mass storage 146. As is well

known in the art, the Client application 12 is loaded and stored in main memory 144 and mass storage 146 and executed by the CPU 142.

5 **[0231]** Referring to Figure 15B, a block diagram 150 illustrating the hardware of a Server 16 (or any of the servers 1202 through 1210) used for storing and executing the server application 78 is shown. The hardware includes a CPU 152, main memory 154, mass storage 156, and the archive 89. As is well known in the art, the server application 78 is loaded and stored in main memory 154 and mass storage 156 and executed by the CPU 152. As noted above, indexed media payloads of one or more Users are stored in the archive 89.

10 **[0232]** Although many of the components and processes are described above in the singular for convenience, it will be appreciated by one of skill in the art that multiple components and repeated processes can also be used to practice the techniques of the system and method described herein. Further, while the invention has been particularly shown and described with reference to specific embodiments thereof, it
15 will be understood by those skilled in the art that changes in the form and details of the disclosed embodiments may be made without departing from the spirit or scope of the invention. For example, embodiments of the invention may be employed with a variety of components and should not be restricted to the ones mentioned above. It is therefore intended that the invention be interpreted to include all variations and
20 equivalents that fall within the true spirit and scope of the invention.

Claims:

1. A method for progressively synchronizing stored copies of voice media transmitted over a network during a conversation that may be selectively conducted in a real-time mode and in a time-shifted mode, comprising:

(a) progressively receiving voice media transmitted from a sending node to a receiving node as the voice media is persistently stored at the sending node, the voice media transmitted with a packet size and packetization interval sufficient to enable the near real-time rendering of the voice media;

(b) progressively and persistently storing the transmitted voice media at the receiving node as the voice media is received;

(c) noting any voice media that was transmitted but not received and persistently stored at the receiving node;

(d) generating requests as needed for any noted voice media;

(e) receiving retransmissions of the noted voice media from the sending node to the receiving node in response to the requests;

(f) repeating (d) and (e) until the sending node and the receiving node each have complete copies of the voice media maintained in persistent storage; and

(g) selectively rendering the transmitted voice media at the receiving node either in the real-time mode as the voice media is received over the network or out of persistent storage in the time-shifted mode.

2. The method of claim 1, further comprising notifying the receiving node of the voice media available for transmission.

3. The method of claim 2, wherein the notification occurs either (i) prior to or (ii) substantially concurrent with the transmission.

4. The method of claim 2, wherein notifying the receiving node of the availability of the voice media available for transmission further comprises one of the following:

(i) receiving the available voice media transmitted from the sending node to the receiving node, whereby the notification occurs by virtue of the receipt of the transmission;

- (ii) when the sending node logs onto the network;
- (iii) each time the receiving node receives voice media from the sending node;
- (iv) periodic notifications received by the receiving node from the sending node;
- (v) the receiving node periodically asks the sending node for the status of voice media to be transmitted from the sending node to the receiving node;
- (vi) the sending node sends to the receiving node a synchronization update message; or
- (vii) any combination of (i) through (vi).

5. The method of any one of the preceding claims, further comprising:
progressively encoding the voice media in an indexed format as the voice media is being created; and

progressively and persistently storing the encoded voice media at the sending node in the indexed format as the encoded voice media is progressively transmitted to the receiving node.

6. The method of any one of the preceding claims, wherein the sending node comprises a communication device, the communication device consisting of one of the following: land-line phone, wireless phone, cellular phone, satellite phone, computer, radio, server, satellite radio, tactical radio or tactical phone.

7. The method of any one of the preceding claims, wherein the receiving node comprises a communication device, the communication device consisting of one of the following: land-line phone, wireless phone, cellular phone, satellite phone, computer, radio, server, satellite radio, tactical radio or tactical phone.

8. The method of any one of the preceding claims 1 to 5, wherein sending node is a server hop on the network, between a node where the voice media was created and the recipient.

9. The method of any one of the preceding claims 1 to 5, wherein the receiving node is a server hop on the network between a node where the voice media was originated created and the recipient.

10. The method of any one of the preceding claims 1 to 5, wherein the receiving node and the sending node are first and second server hops on the network respectively.
11. The method of any one of the preceding claims 1 to 5, wherein either the sending node or the receiving node is a server on the network.
12. The method of any one of the preceding claims , wherein noting any voice media that was transmitted but not received and persistently stored further comprises:
 - checking the voice media received at the receiving node from the sending node;
 - noting missing, corrupted or reduced bit rate representations of the voice media as the media is received at the receiving node;
 - maintaining a data store at the receiving node for the noted voice media; and
 - removing the noted voice media from the data store as the noted media is received at the receiving node.
13. The method of claim 12, wherein generating requests as needed for any noted media that is not already locally stored at the receiving node further comprises:
 - scanning the data quality store; generating receipt reports, the receipt reports including requests for retransmission of the voice media noted in the data store; and
 - transmitting the receipt reports from the receiving node to the sending node.
14. The method any one of the preceding claims , further comprising receiving and persistently storing at the receiving node, in addition to the voice media, one or more of the following types of media: video, text, sensor data, position or GPS information, radio signals, or a combination thereof.
15. The method of any one of the preceding claims , wherein the persistently stored voice media at the sending node and the receiving node is synchronized when the sending node and the receiving node each have complete copies of the voice media.
16. The method of any one of the preceding claims , wherein persistently storing means at least one of the following:
 - (i) storage at the receiving node beyond what is needed to render the voice media in real-time as the voice media is received;

(ii) storage at the sending beyond what is needed to transmitted the voice media in real-time as the voice media is transmitted; or (iii) both (i) and (ii).

17. The method of any one of the preceding claims , wherein the voice media is synchronized when the sending node and the receiving node each have persistently stored complete copies of the voice media as originally encoded when created.

18. A programmable communication device configured to carry out the method of any one of the claims 1 through 17.

19. A computer readable medium bearing an application executable by a data processor of a communication device for configuring the communication device to carry out the method according to any one of claims 1 through 17.

20. A method performed on a client communication device, comprising:

synchronizing copies of persistently stored transmitted and received voice messages exchanged between the client communication device and a remote client communication device, the synchronization performed until both the client communication device and the remote client communication device have persistently stored complete copies of both the voice messages transmitted by the client communication device to the remote client communication device and the voice messages received by the client communication device from the remote client communication device, wherein the synchronization occurs and completes, if necessary, beyond a time when it is not possible to render at least one of the transmitted and received voice messages in real-time; and

selectively rendering the voice media of the received messages on the client communication device in both:

(i) a real-time mode as the voice media is received from the remote client communication device; and

(ii) a time-shifted mode by retrieving and rendering the received voice media from persistent storage,

wherein the synchronization occurs and completes, as necessary, beyond a time at which it is no longer possible to render the voice media in real-time, and

wherein persistently stored means for a time period beyond what is needed for the client communication device to:

transmit the voice media of the transmitted messages; and
render the received voice media of the received voice messages in the real-time mode.

21. The method of claim 20, further comprising selectively transitioning the rendering of the voice media of the received messages between the modes (i) and (ii).

22. The method of claim 20 or 21, wherein the voice media of the transmitted and received messages is live voice media that is:

(i) transmitted by the client communication device as the voice media is created; or

(ii) received as the voice media is transmitted by the remote client communication device.

23. The method of any one of the preceding claims 20 to 22, wherein the transmitted and received voice media of the messages is streamed over a network as the voice media is created.

24. The method of any one of the preceding claims 20 to 23, further comprising providing one or more of the following rendering options for rendering the voice media of the received messages on the client communication device:

(i) play;

(ii) play faster;

(iii) play slower;

(iv) jump forward;

(v) jump backward;

(vi) jump to the most current media; and

(vii) pause; and

(viii) catch up to the most recently received voice media of an incoming message.

25. The method of any one of the preceding claims 20 to 24, further comprising persistently storing at a storage location accessible by the client communication device the complete copies of the voice media of the transmitted and received messages.

26. The method of any one of the preceding claims 20 to 24, further comprising persistently storing the voice media of the transmitted and received messages in a time-indexed format.

27. The method of any one of the preceding claims 20 to 26, further comprising associating selected transmitted and received voice messages with one or more conversations.

28. The method of any one of the preceding claims 20 to 27, further comprising noting voice media missing from the voice messages received from the remote client communication device.

29. The method of claim 28, generating requests for retransmission of the voice media noted as missing.

30. The method of any one of the preceding claims 20 to 29, further comprising:
noting missing, corrupted or reduced bit rate representations of the voice media of the received messages;
maintaining a data quality store for the noted voice media; and
removing the noted voice media from the data quality store as the noted voice media is received at the client communication device as the requests for retransmission are satisfied.

31. The method of claim 30, wherein the requests for retransmission are generated by:
scanning the data quality store;
generating receipt reports, the receipt reports including the requests for retransmission of the voice media noted in the data quality store; and
transmitting the receipt reports from the client communication device to the remote client communication device.

32. The method of any one of the preceding claims 20 to 31, further comprising synchronizing the transmitted voice messages between the client communication device and the remote client communication device by retransmitting voice media noted as missing by the remote client communication device so that both devices have persistently stored complete copies of the transmitted voice messages.

33. The method of any one of the preceding claims 20 to 32, further comprising progressively transmitting the voice media of a select message from the client communication device to the remote client communication device by:

- (i) defining a transmission loop;
- (ii) determining the voice media of the select message available for transmission during the defined transmission loop;
- (iii) ascertaining bandwidth on a network coupling the client communication device and the remote client communication device during the transmission loop;
- (iv) comparing the ascertained bandwidth on the network with the bandwidth necessary to transmit the available voice media; and
- (v) transmitting the available voice media if the ascertained bandwidth is sufficient during the transmission loop.

34. The method of claim 33, further comprising progressively transmitting the available voice media by:

- (vi) determining if the ascertained bandwidth is not sufficient for transmitting the available voice media during the transmission loop;
- (vii) generating a reduced bit rate representation of the available voice media; and
- (viii) transmitting the reduced bit rate representation when the available bandwidth on the network is not sufficient during the transmission loop,

whereby the transmission of the reduced bit rate representation enhances the timeliness of the delivery of the voice media so as to increase the ability of the remote client communication device to render the media upon receipt.

35. The method of claim 34, further comprising generating the reduced bit rate representation of the available voice media by performing one or more of the following:

- (a) using fewer bits per media unit of time when packetizing the reduced bit rate representation of the media relative to a bit rate representation of the media as originally encoded;
- (b) adjusting the packetization interval for packets used to packetized the reduced bit rate representation of the media;
- (c) adjusting the rate of transmission of the packets used to packetized the reduced bit rate representation of the media; or
- (d) any combination of (a) through (c).

36. The method of claim 35, further comprising, for the duration of the select message, continually defining transmission loops and performing (i) through (viii) and (a) through (c) as needed for each transmission loop.

37. The method of any one of the preceding claims 20 to 36, further comprising transmitting and receiving messages containing one or more of the following types of media: video, voice, text, sensor data, position or GPS information, radio signals, or a combination thereof.

38. The method of any one of the preceding claims 20 to 37, wherein the client communication device performing the method comprises one of the following: land-line phone, wireless phone, cellular phone, mobile phone, satellite phone, computer, radio, server, satellite radio, tactical radio or tactical phone.

39. A programmable communication device configured to carry out the method of any one of the claims 20 through 38.

40. A computer readable medium bearing an application executable by a data processor of a communication device for configuring the communication device to carry out the method according to any one of claims 20 through 38.

41. A method for progressively synchronizing stored copies of voice media transmitted over a network during a conversation that may be selectively conducted in a real-time mode and in a time-shifted mode; and/or

a programmable communication device; and/or

a computer readable medium bearing an application executable by a data processor of a communication device; and/or

a method performed on a client communication device substantially as herein described with reference to any one of the embodiments of the invention illustrated in the accompanying drawings and/or examples.

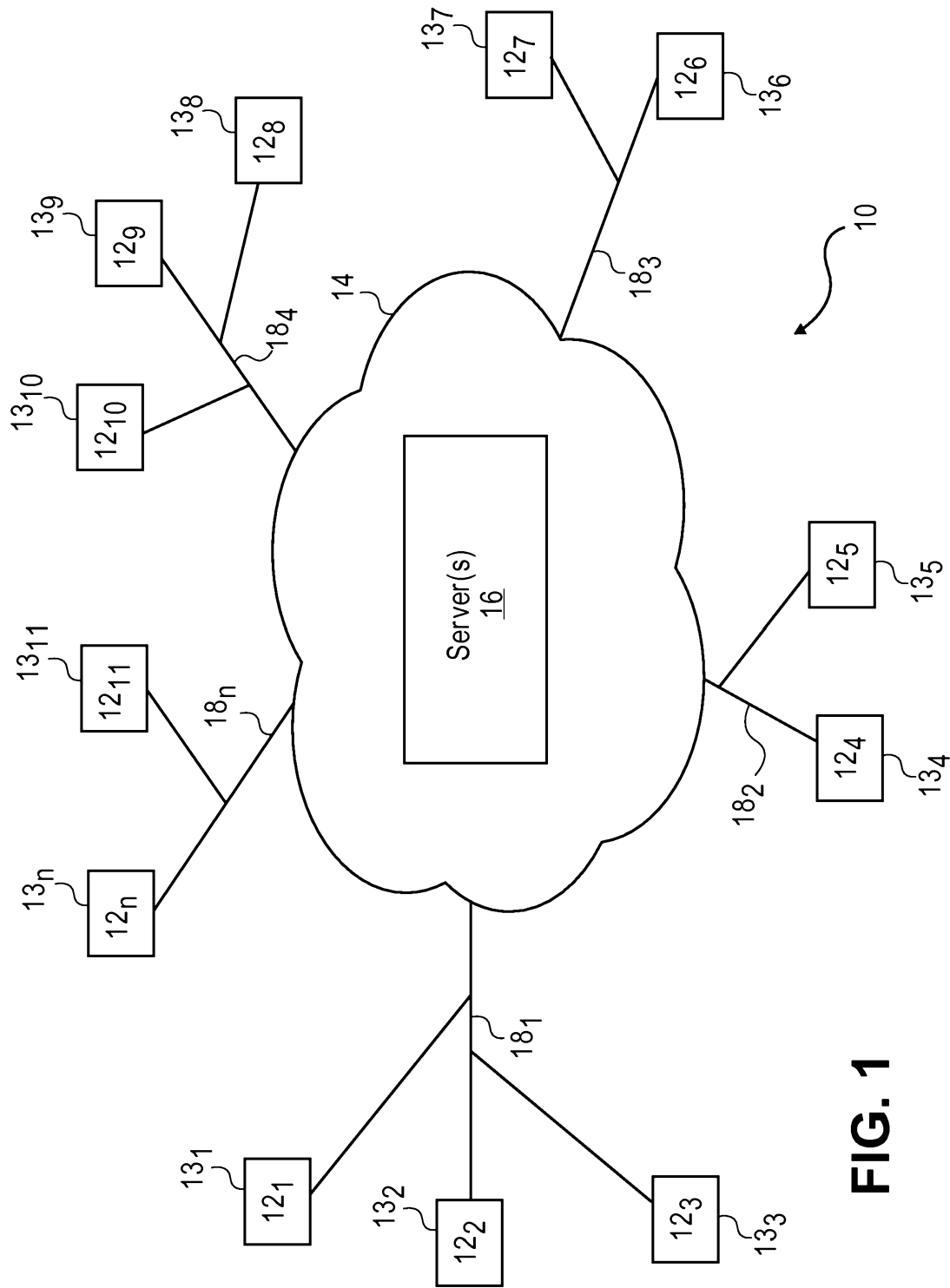


FIG. 1

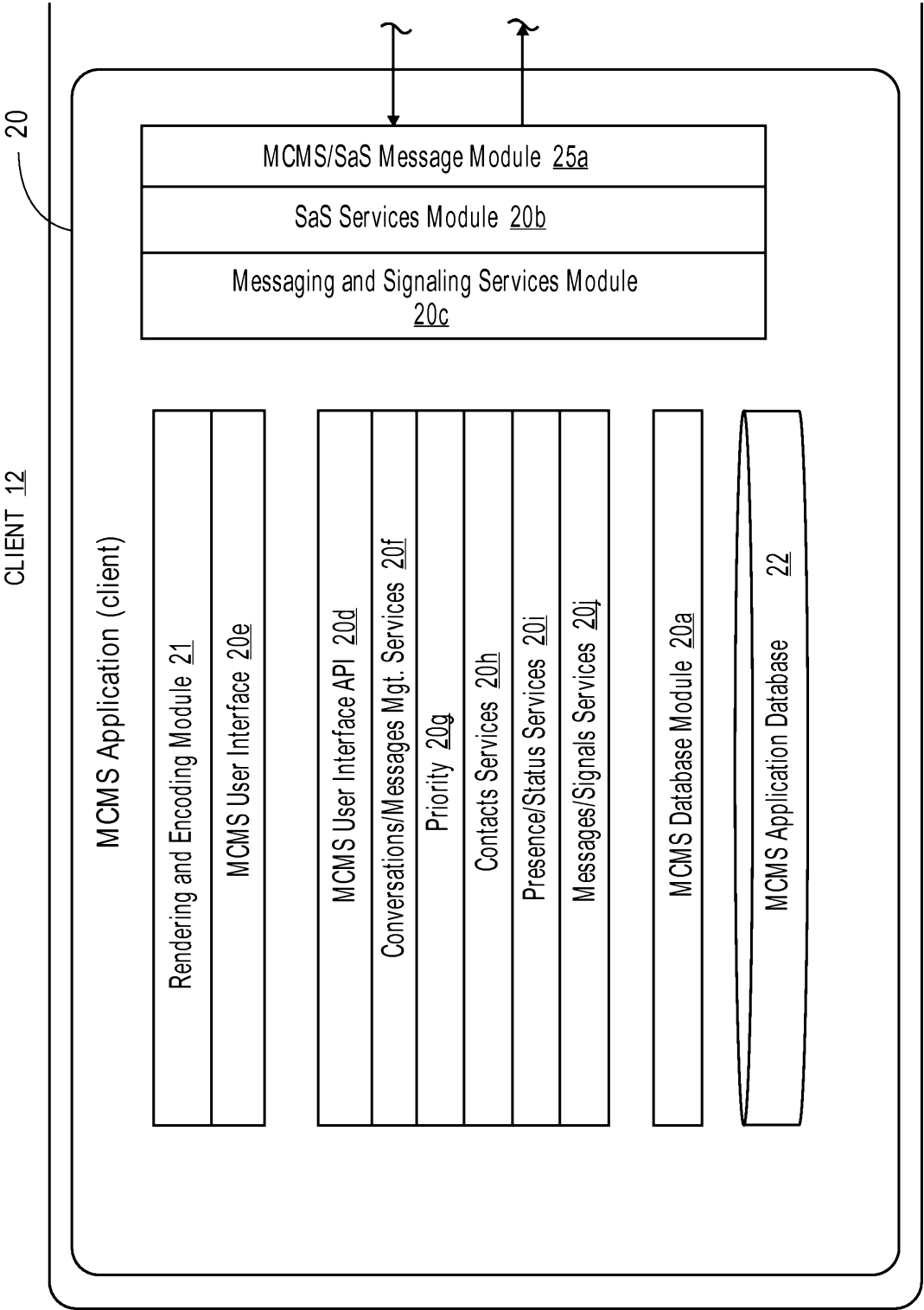


FIG. 2A FIG. 2B

FIG. 2A

CLIENT 12

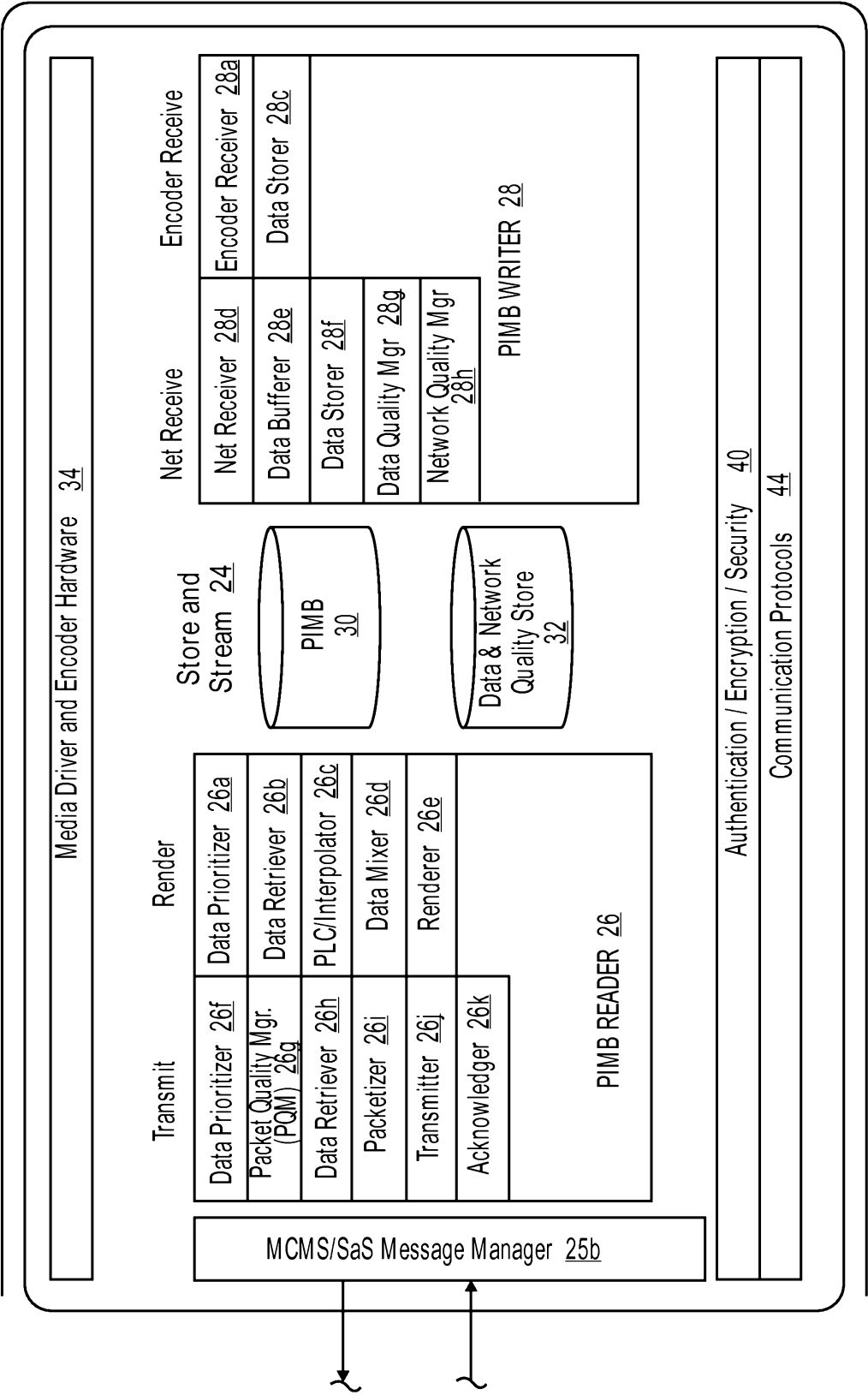


FIG. 2B

FIG. 2A FIG. 2B

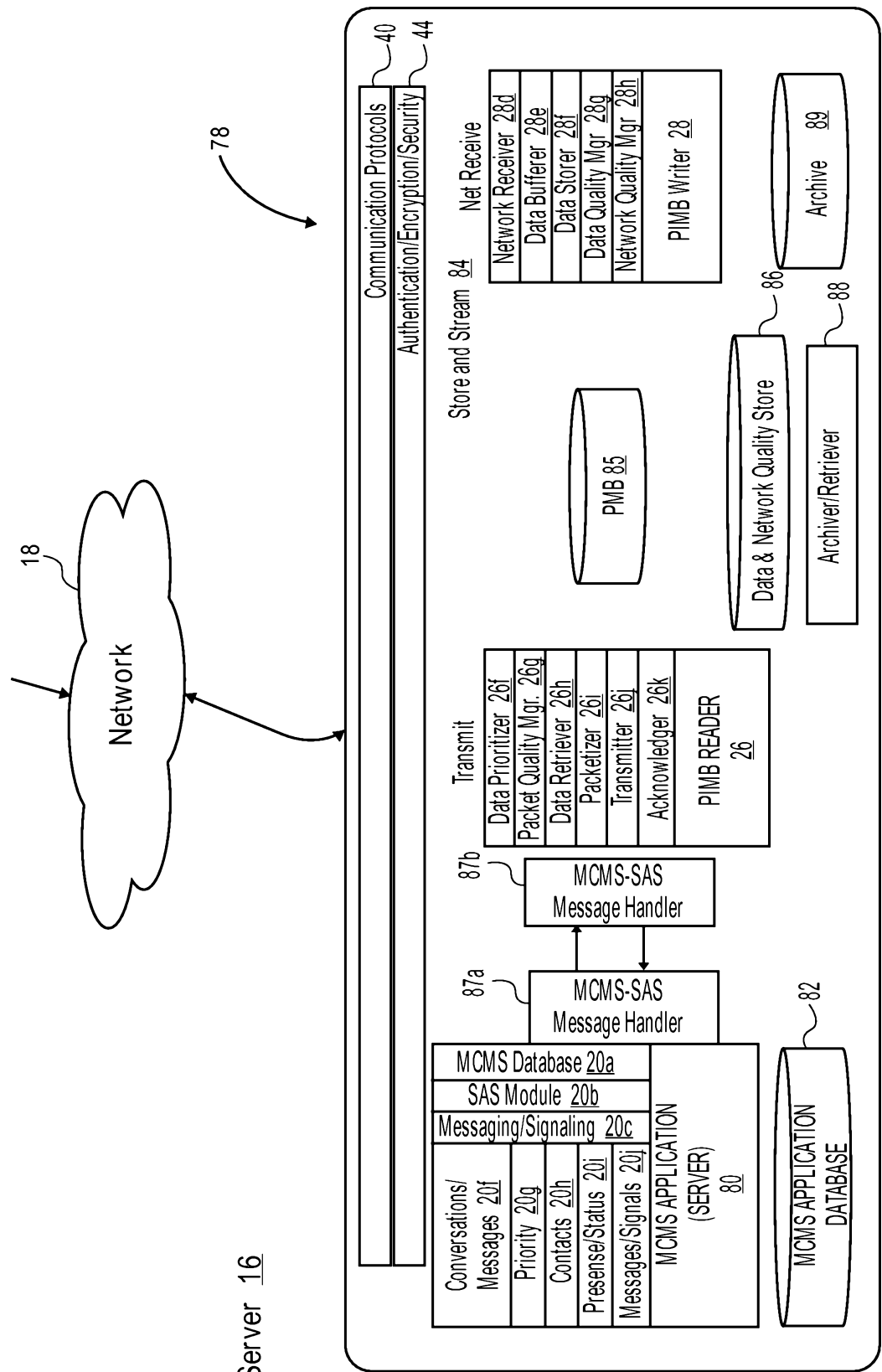


FIG. 3

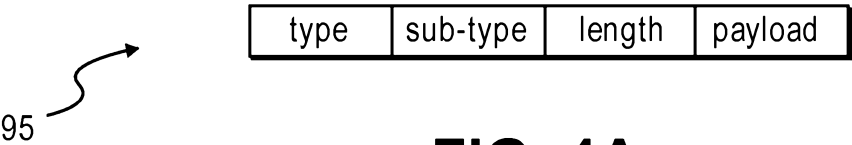


FIG. 4A

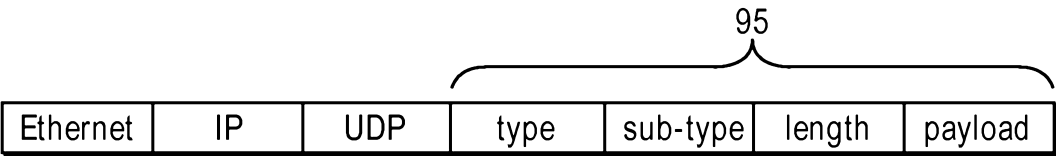
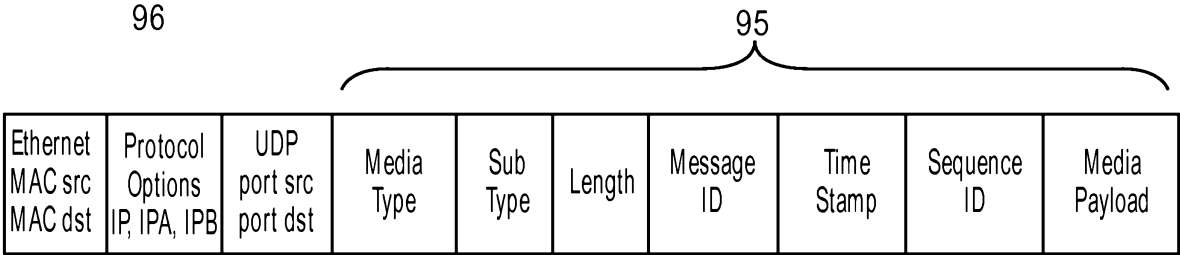
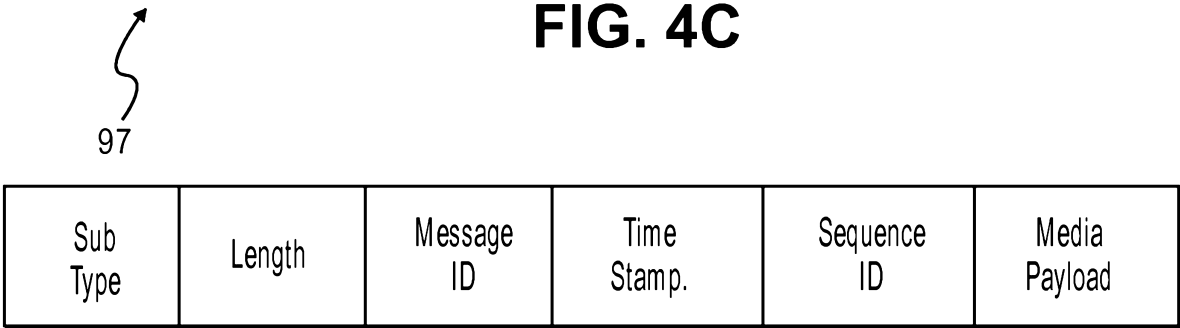


FIG. 4B



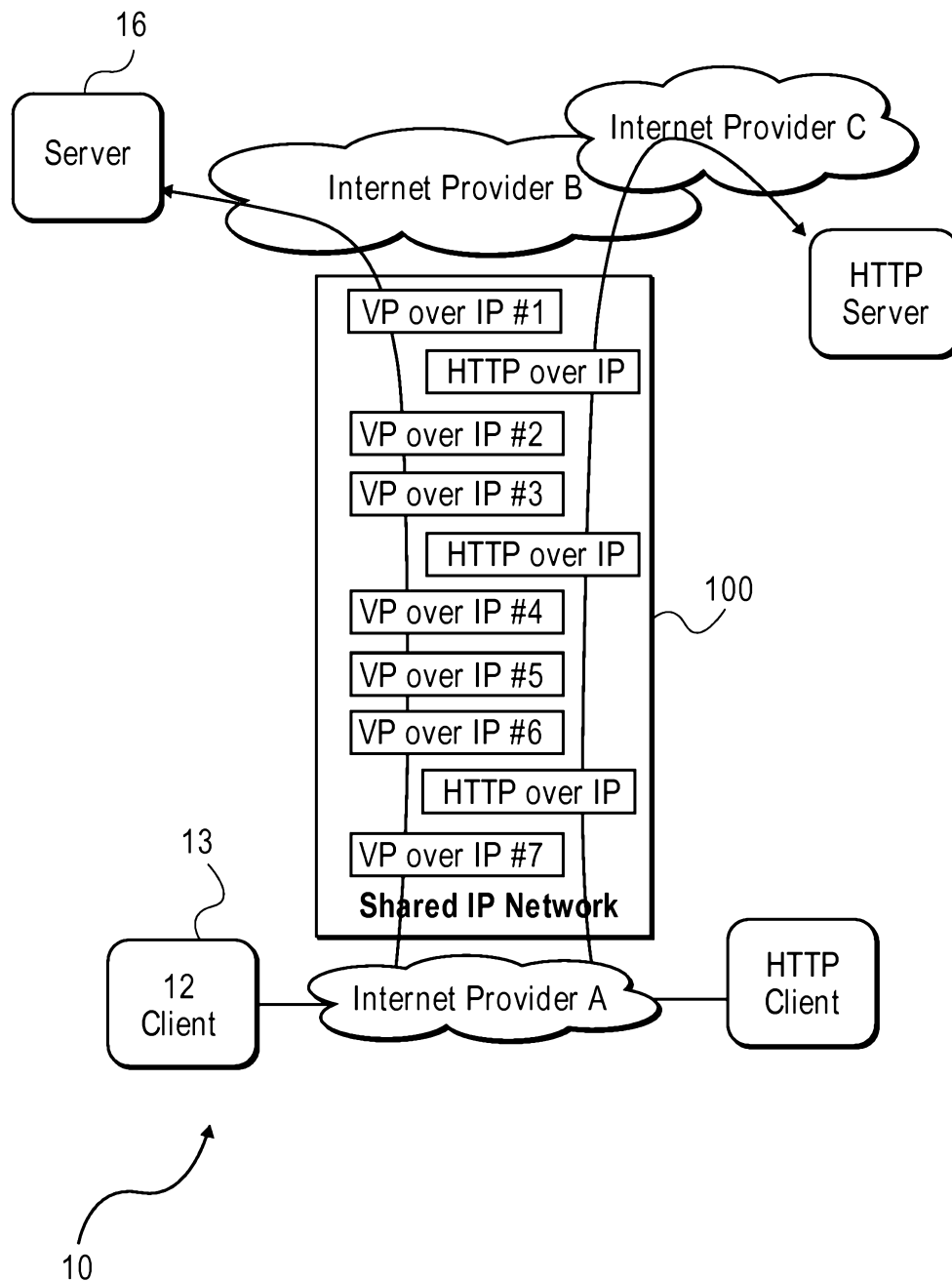
96

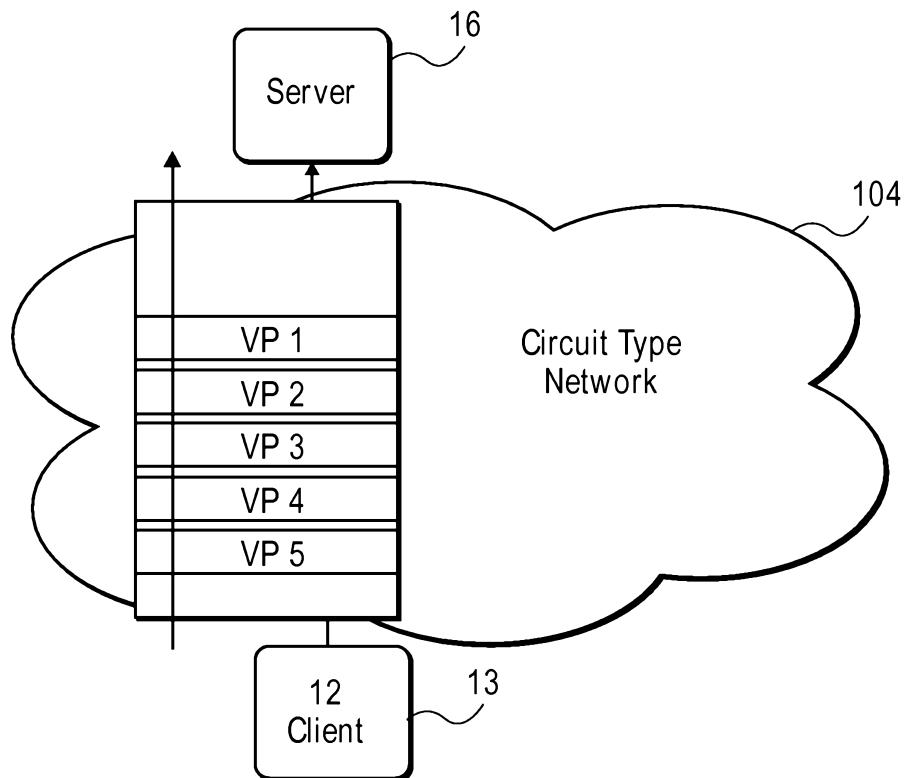
FIG. 4C



98

FIG. 4D

**FIG. 5**

**FIG. 6**

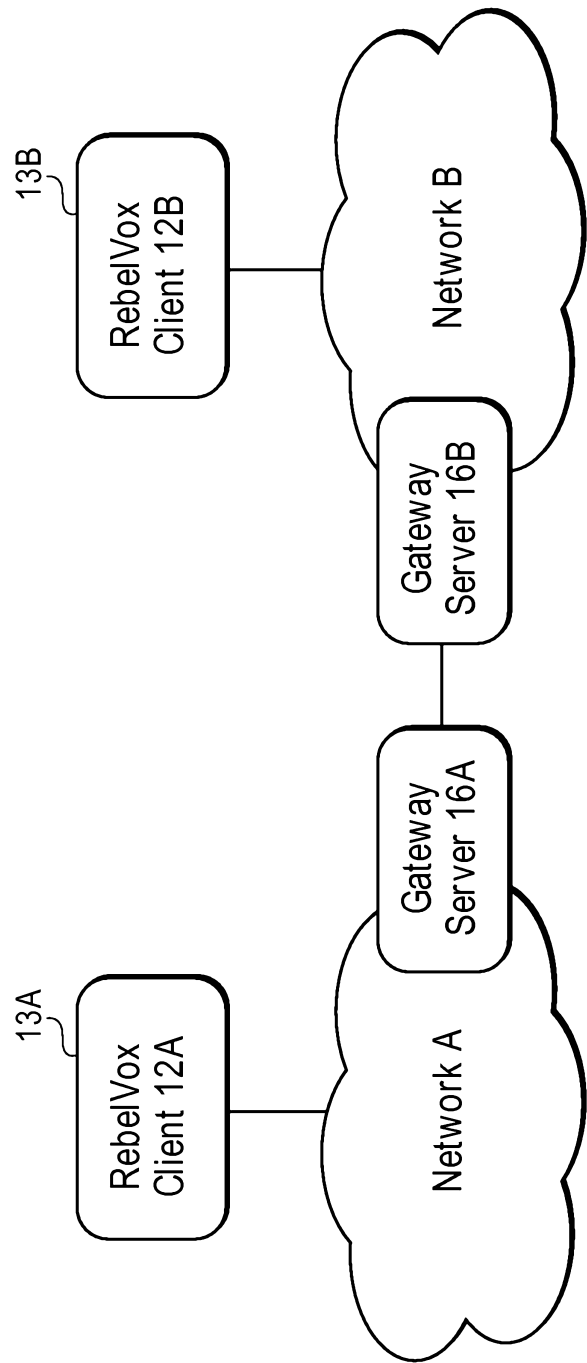


FIG. 7

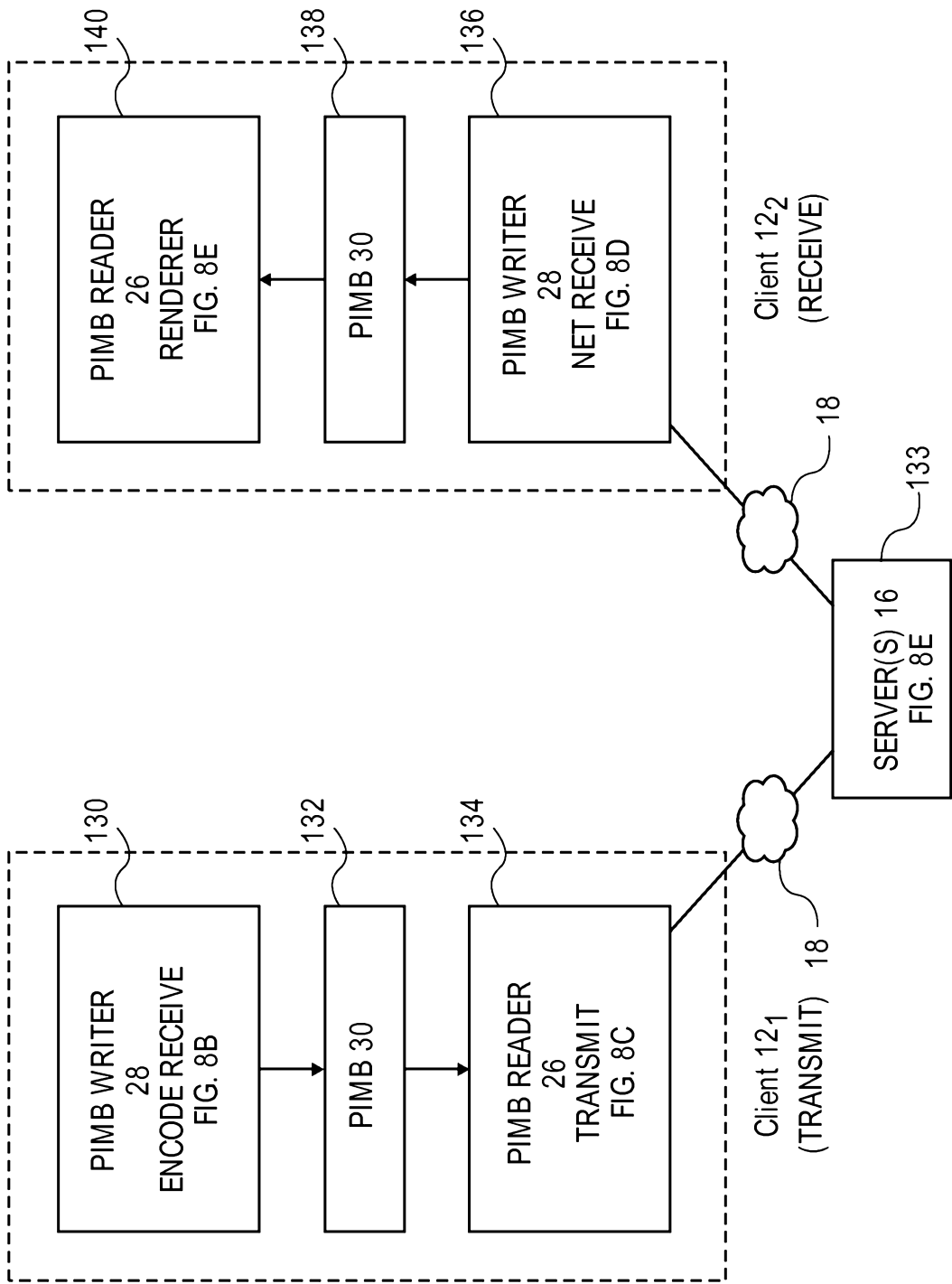
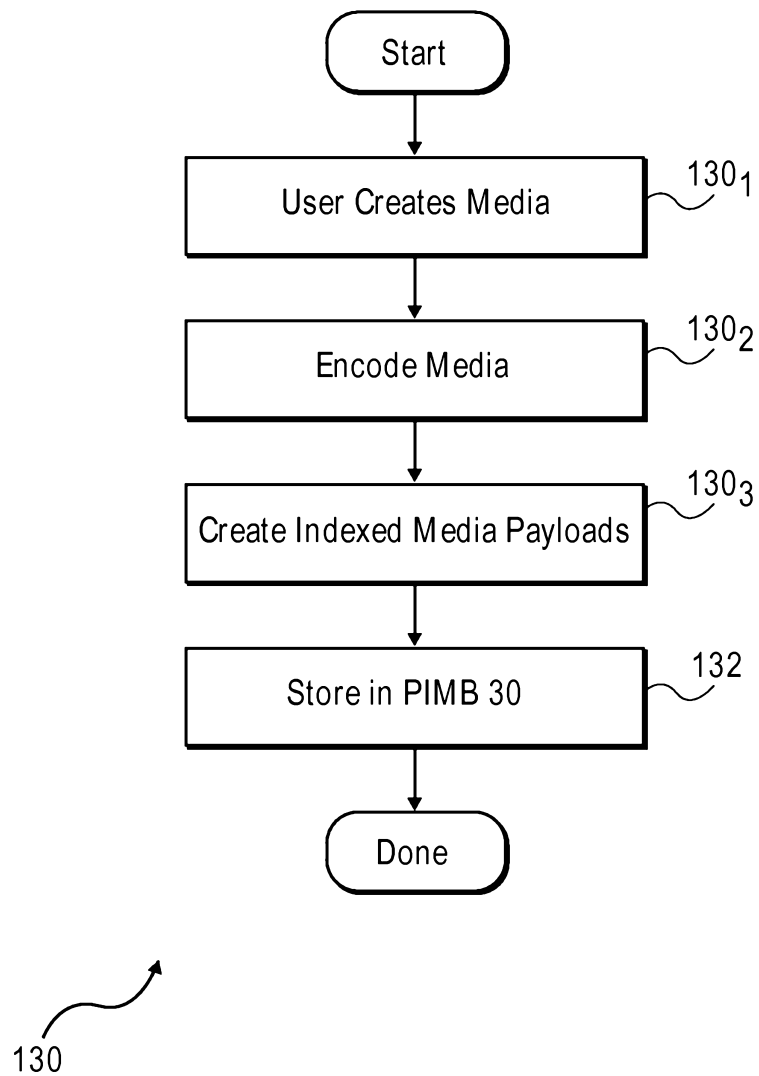
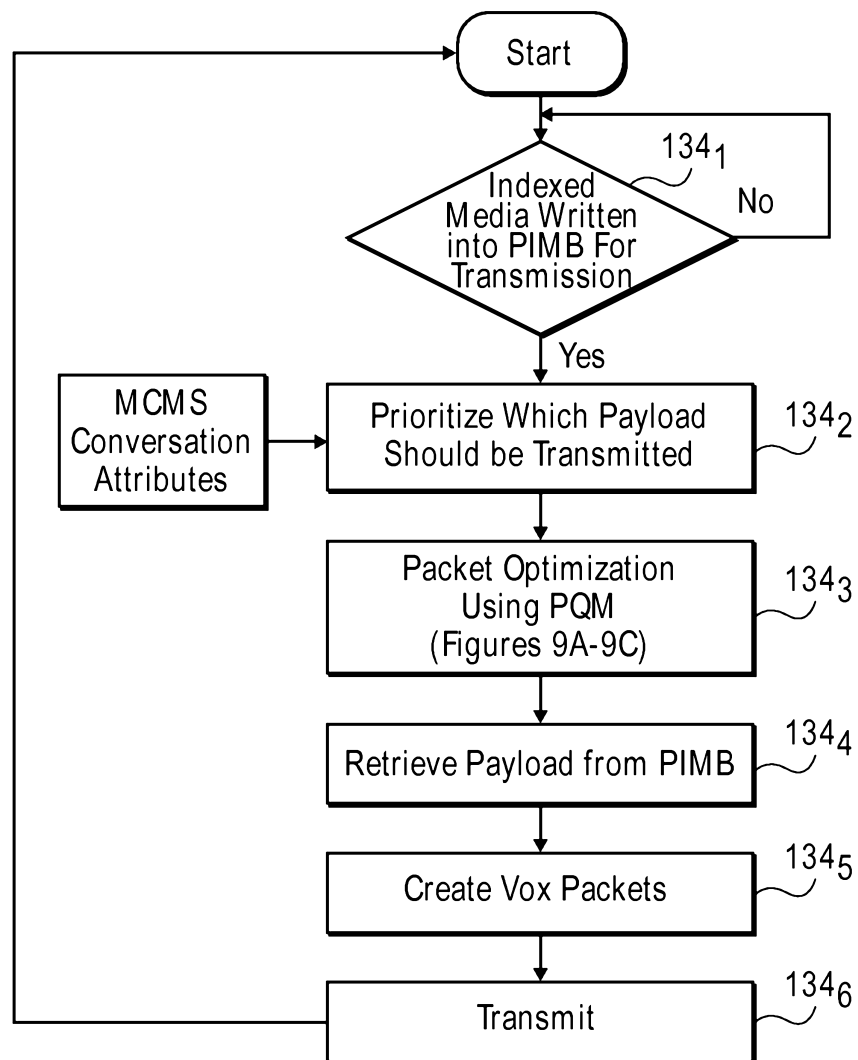
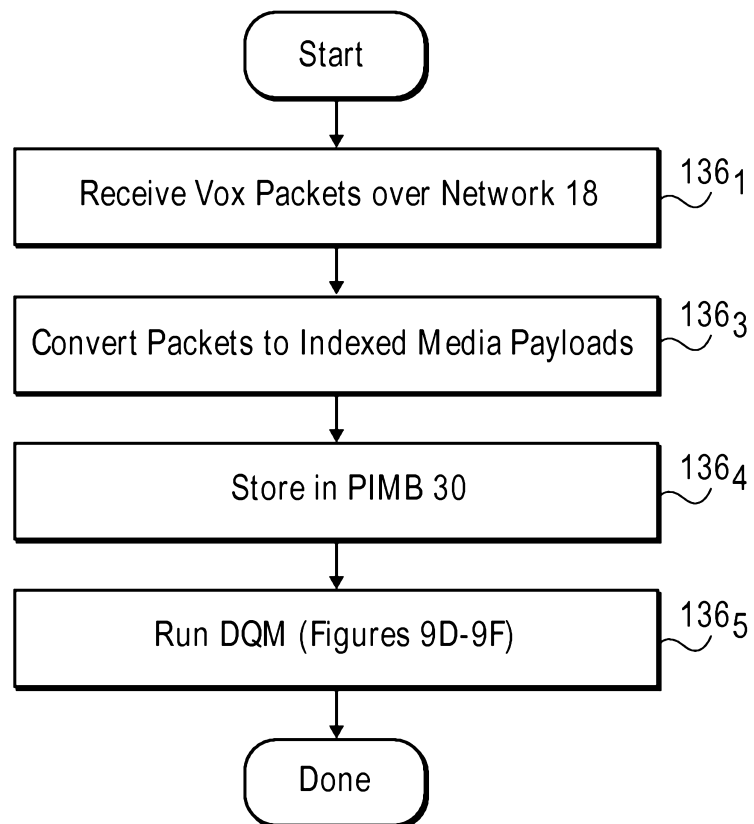


FIG. 8A

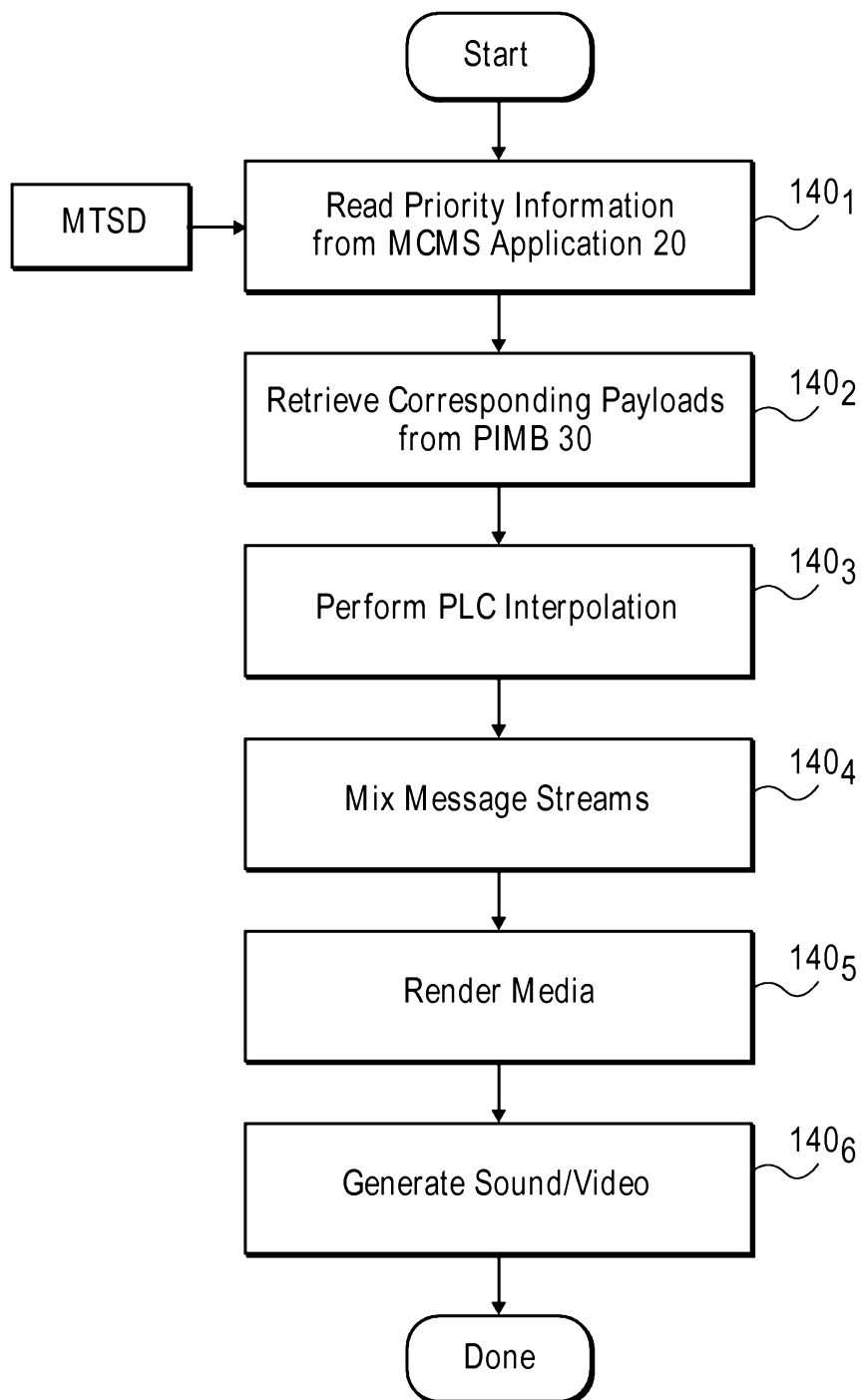
**FIG. 8B**

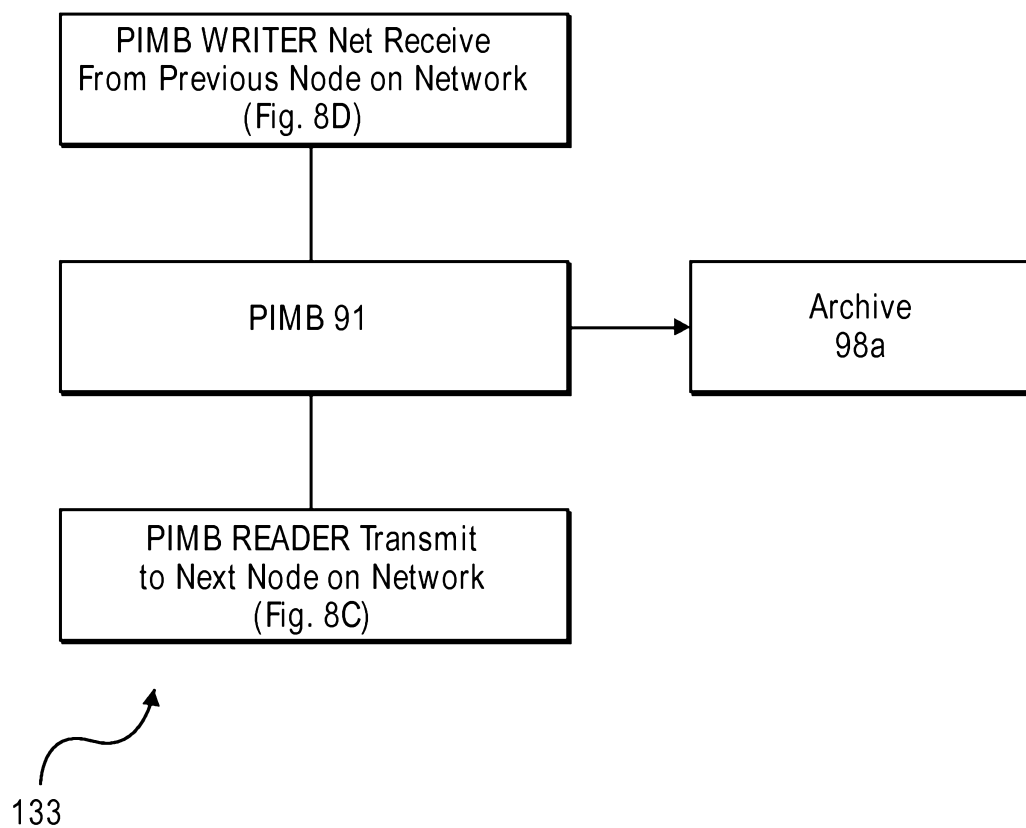
**FIG. 8C**

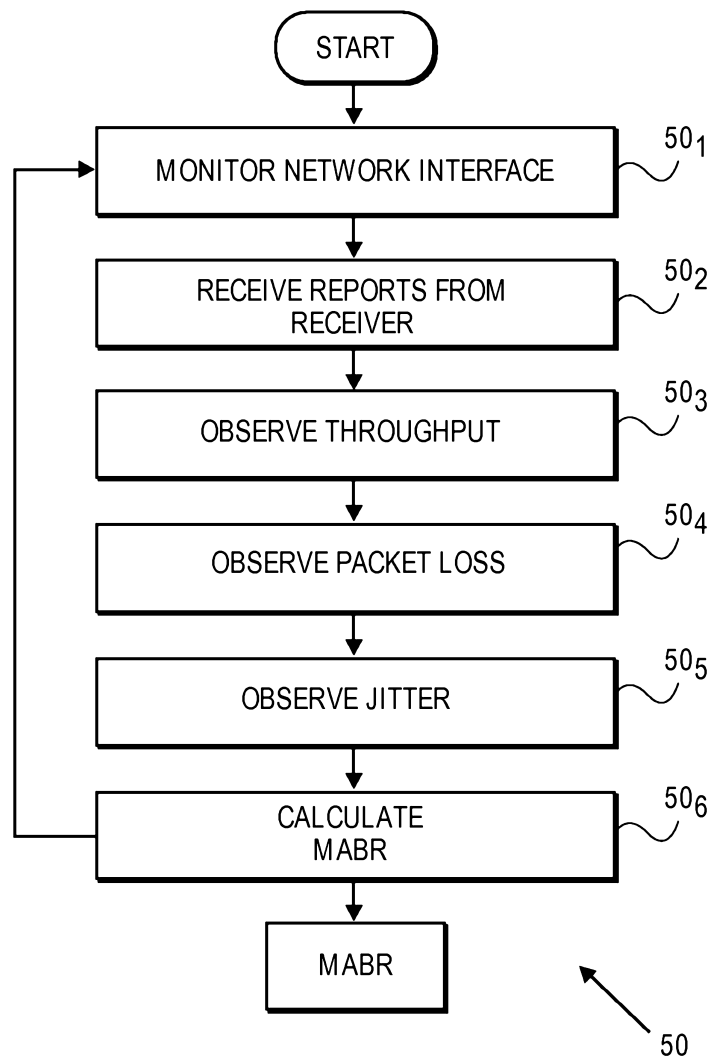
12 of 26

**FIG. 8D**

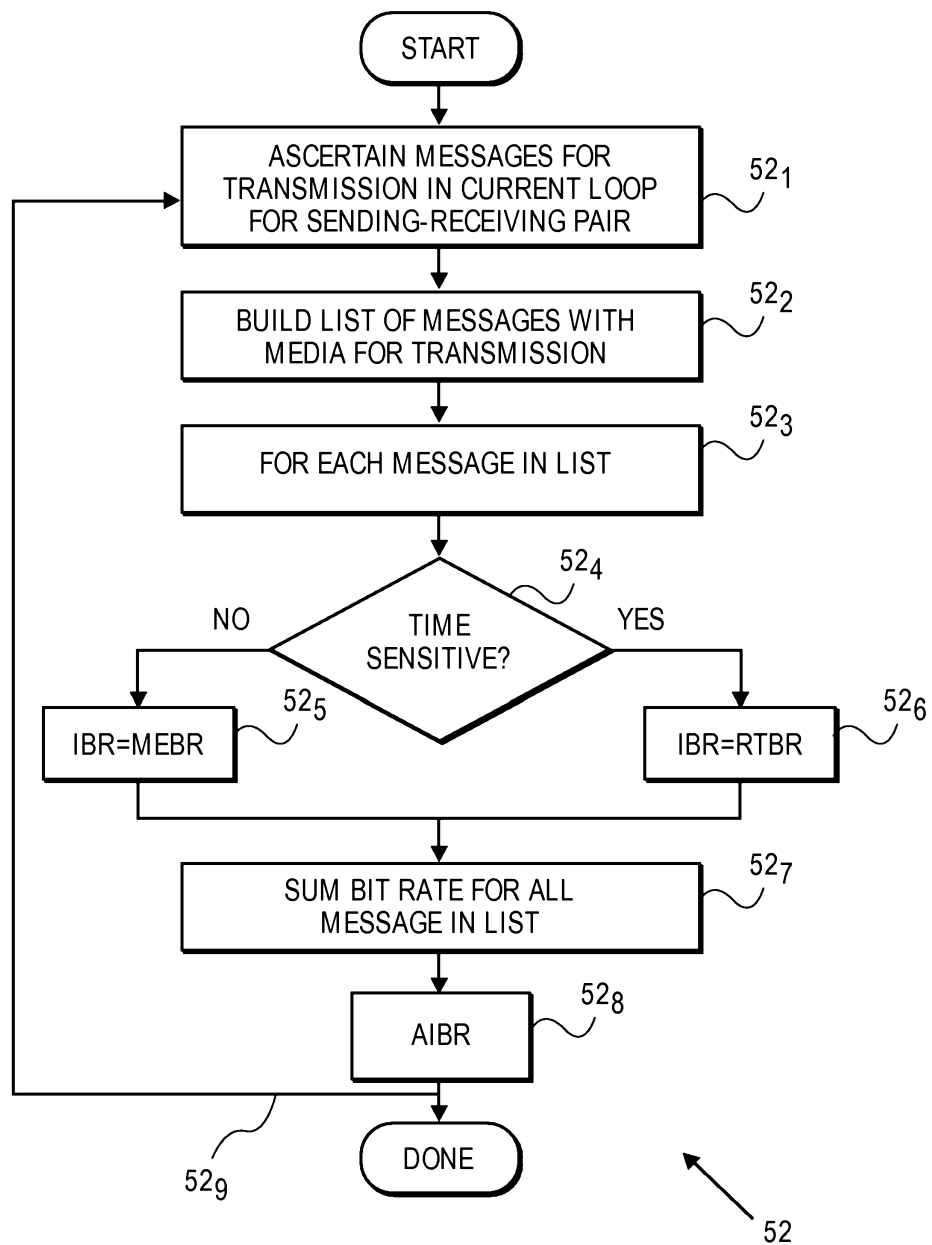
13 of 26

**FIG. 8E**

**FIG. 8F**

**FIG. 9A**

16 of 26

**FIG. 9B**

17 of 26

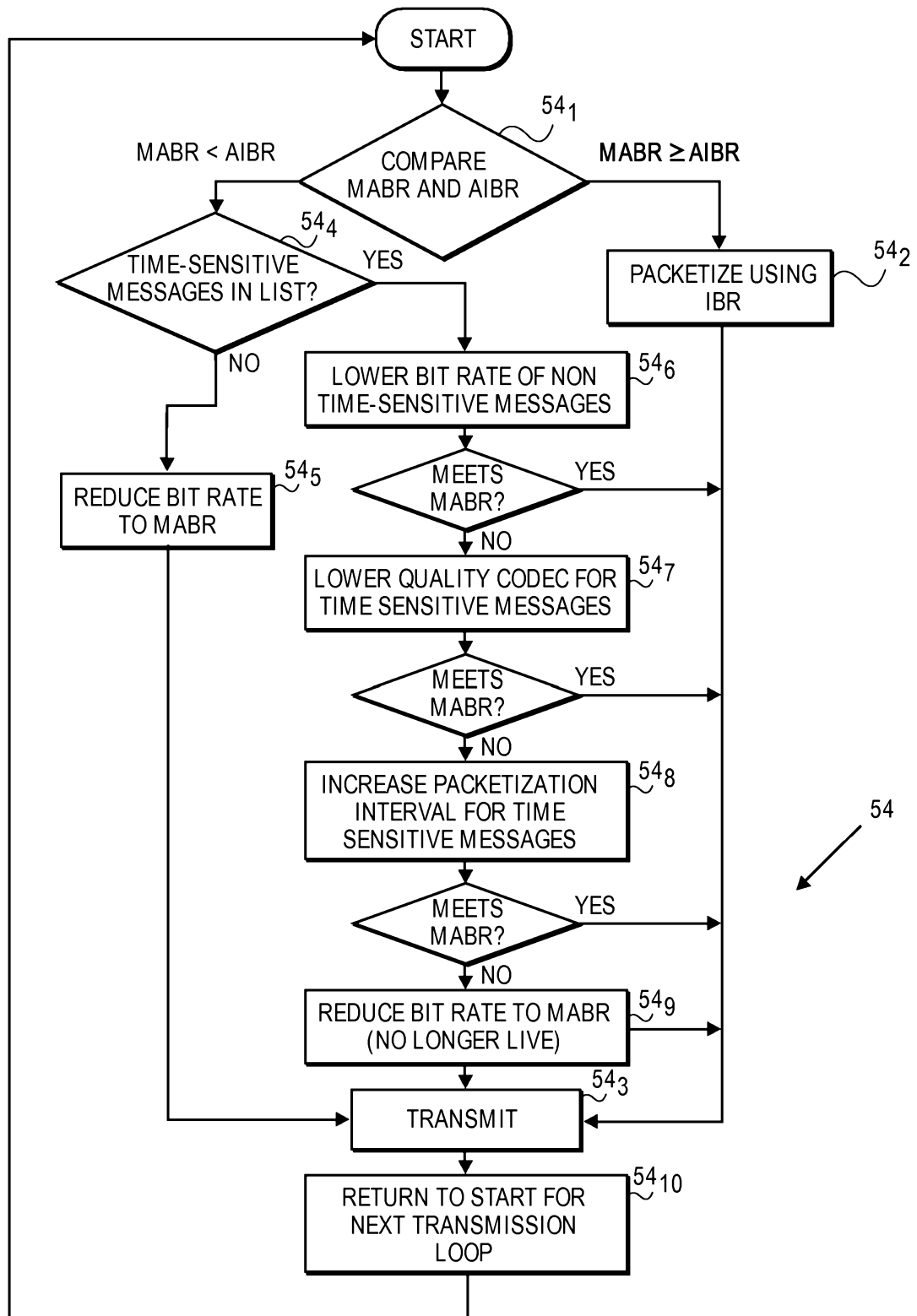
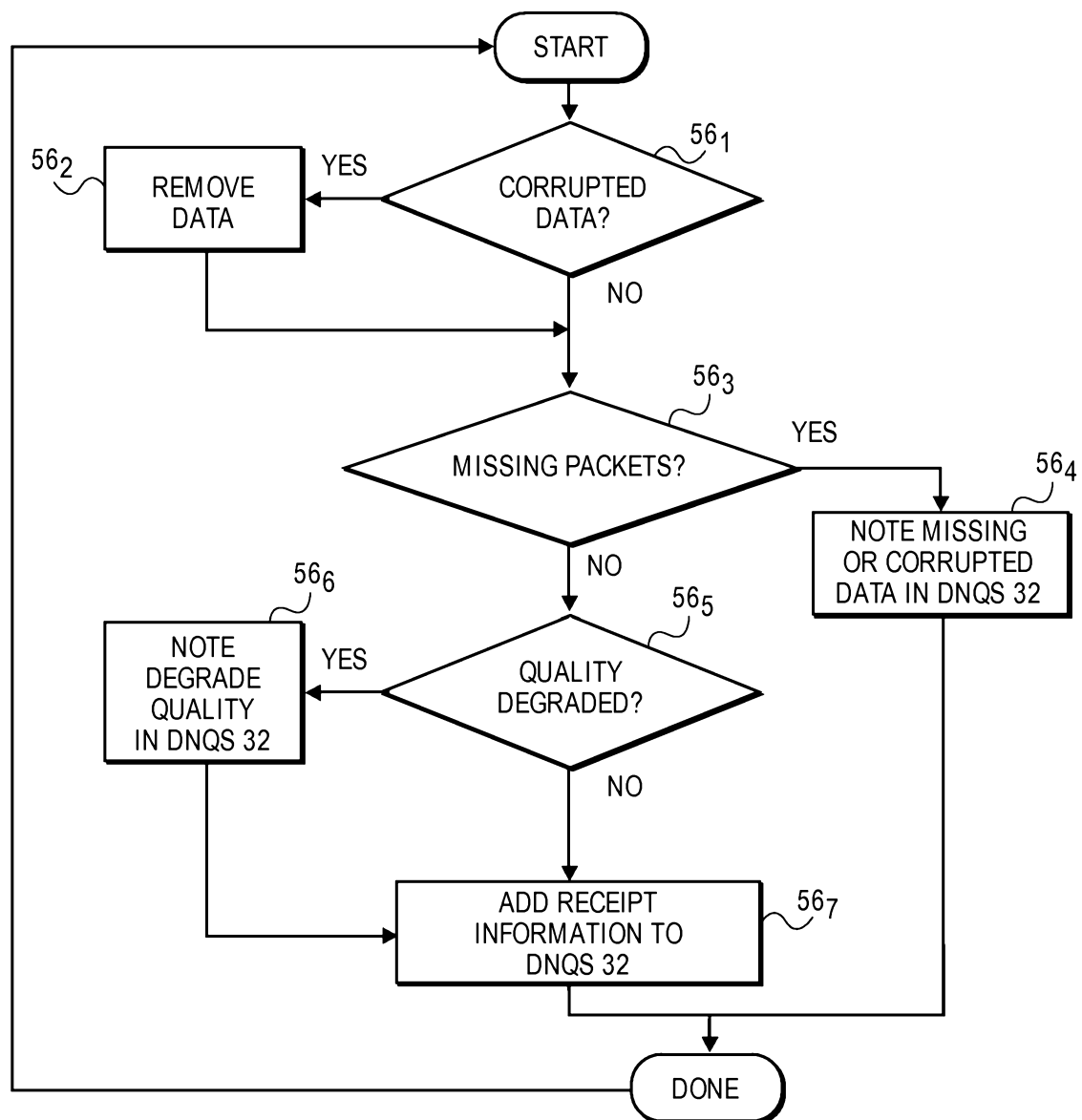
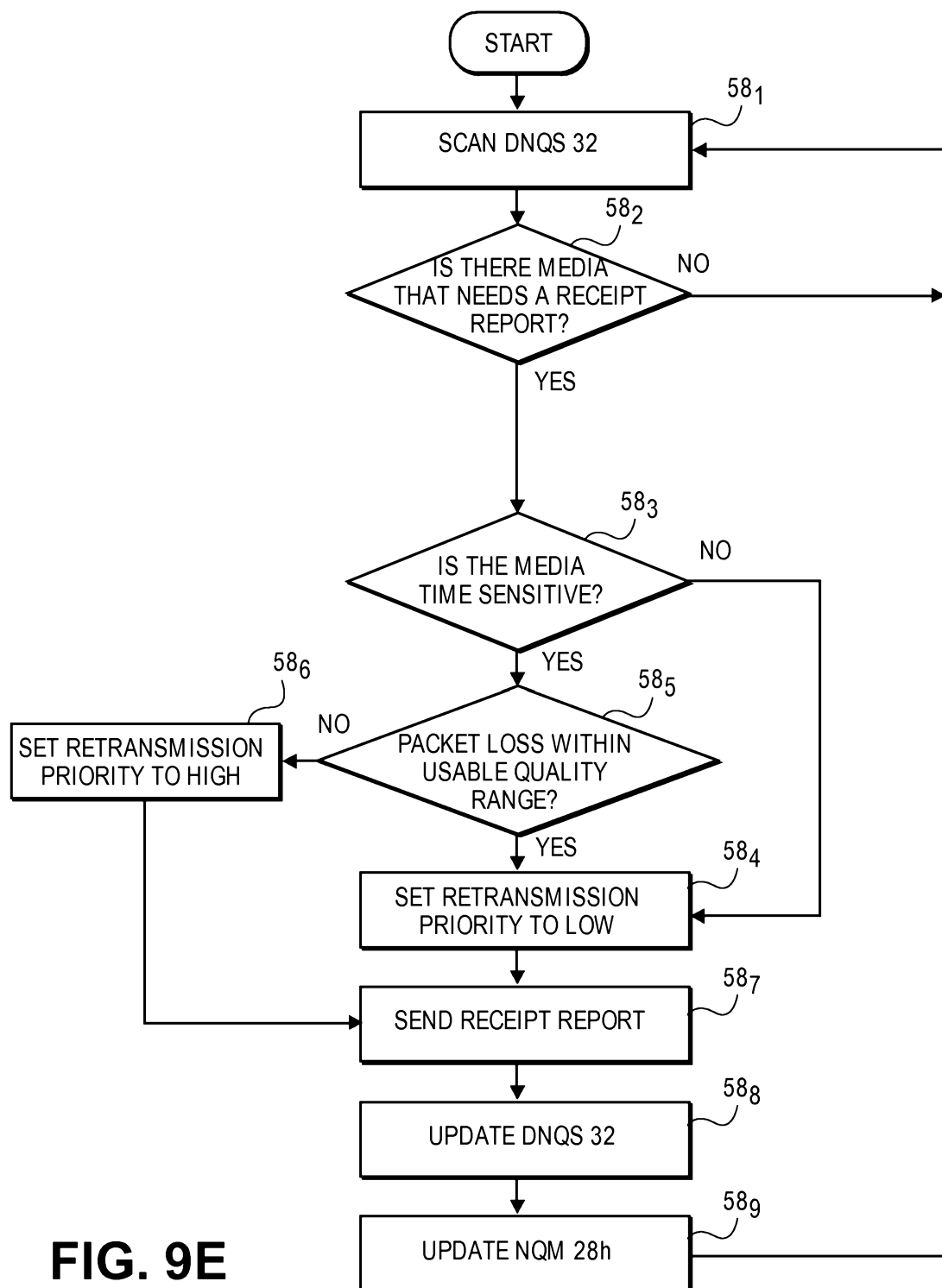


FIG. 9C

**FIG. 9D**

19 of 26



20 of 26

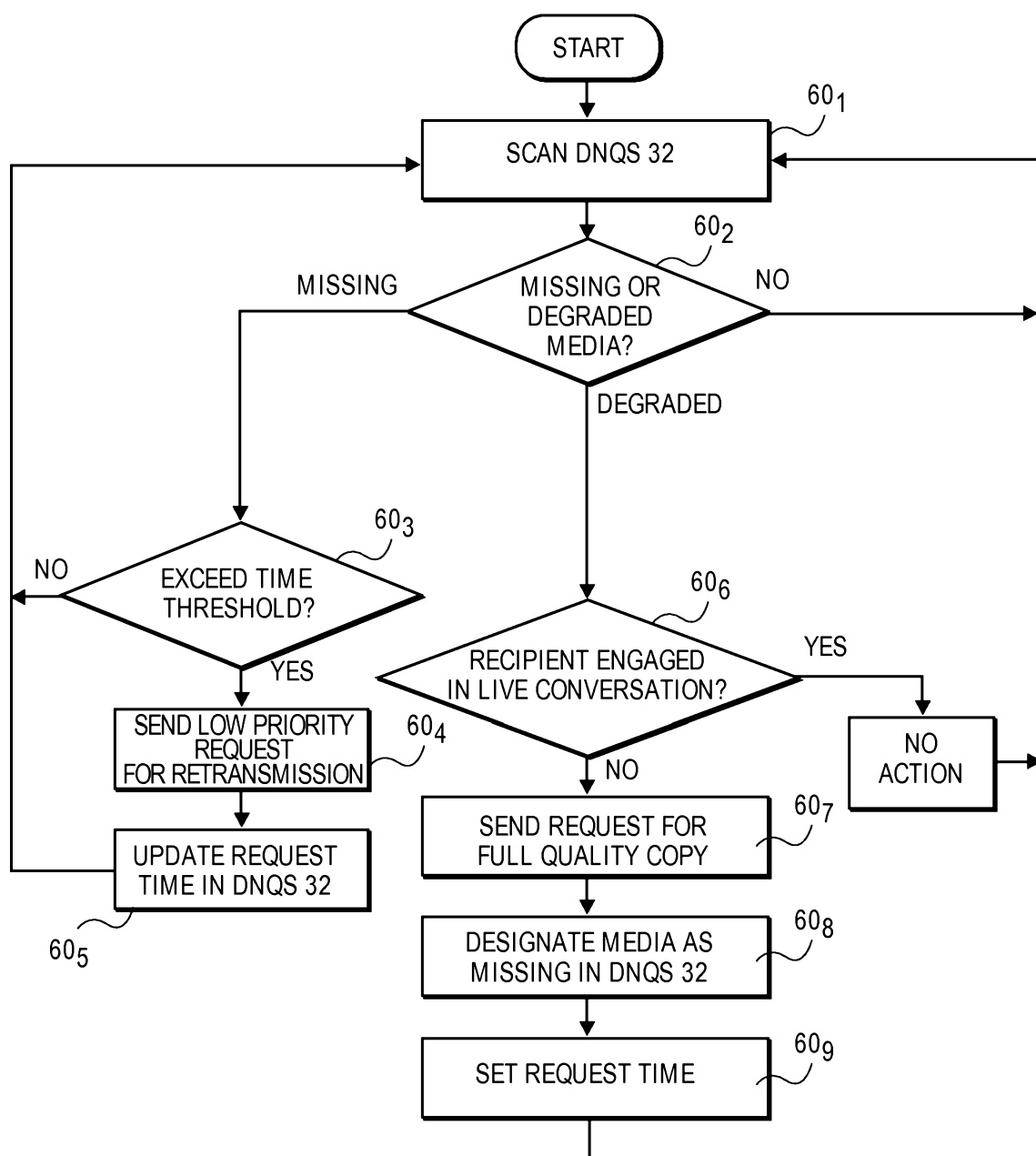


FIG. 9F

Real Time Synchronization for
Voice Communications

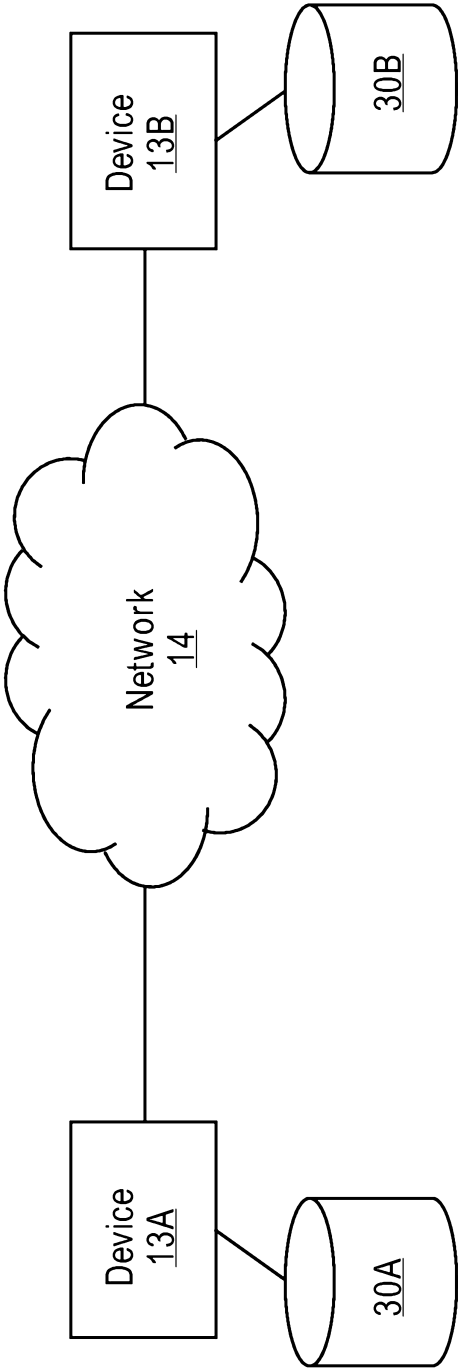
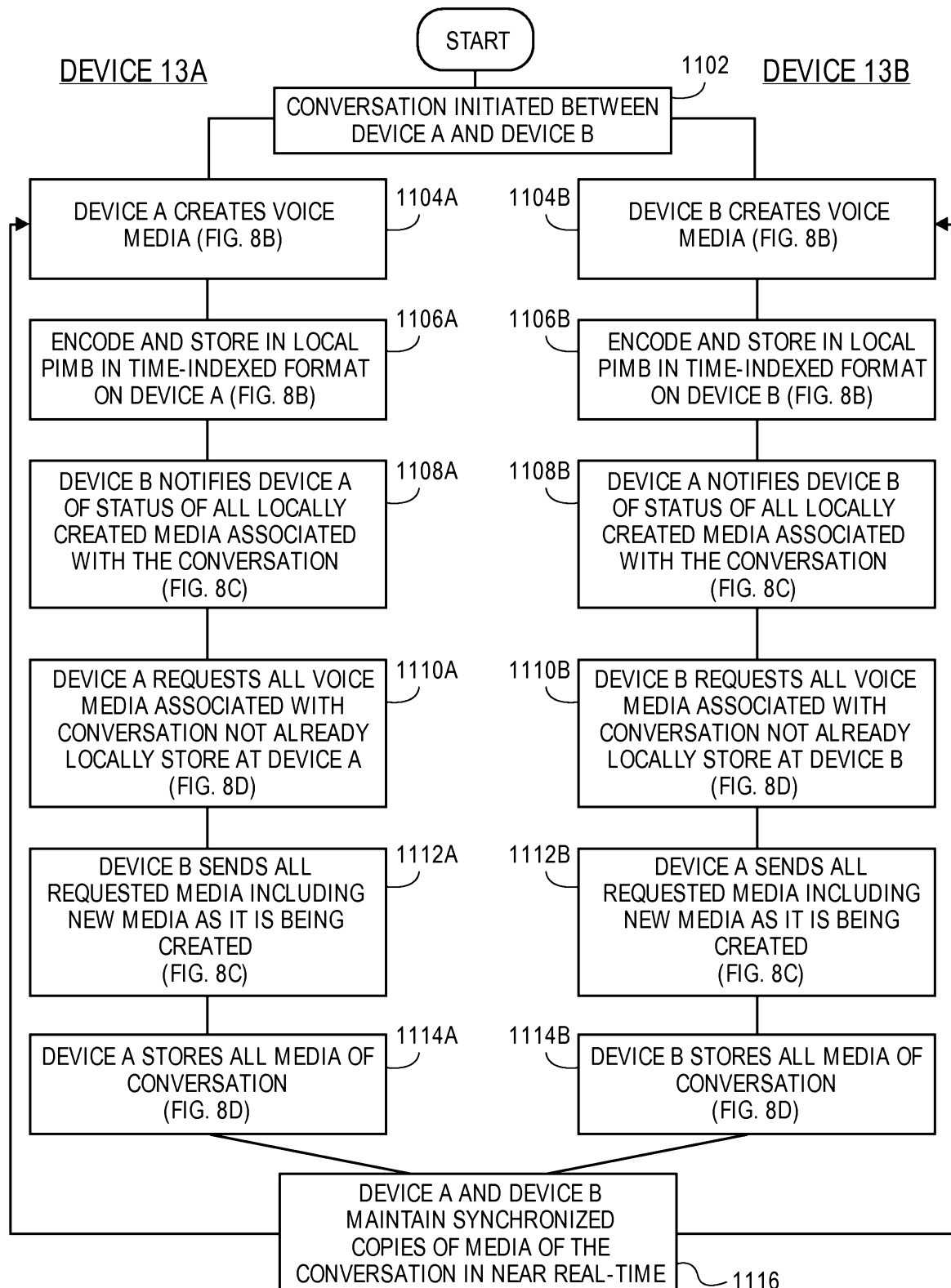
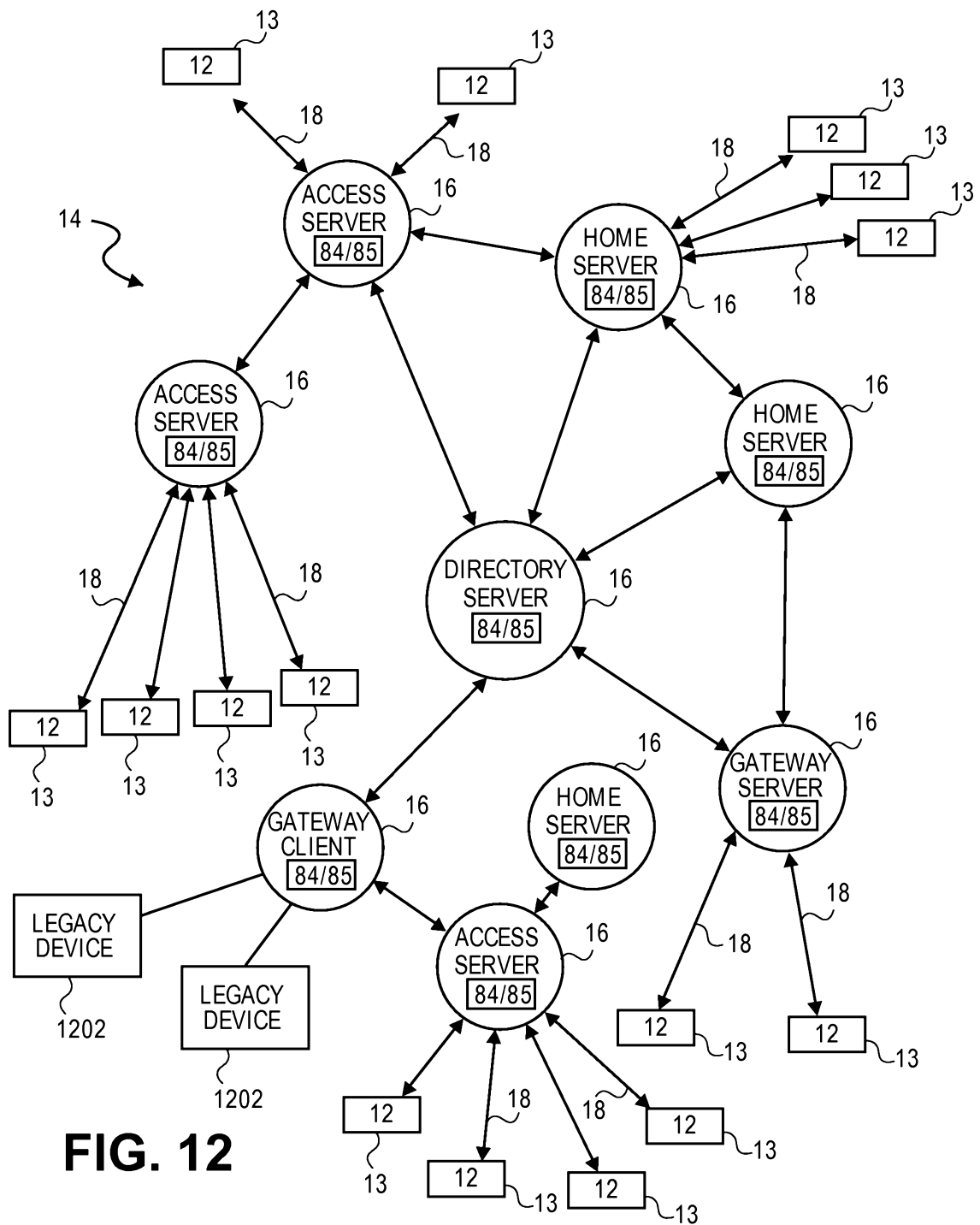


FIG. 10

**FIG. 11**

Distributed Services Architecture

Network Routing

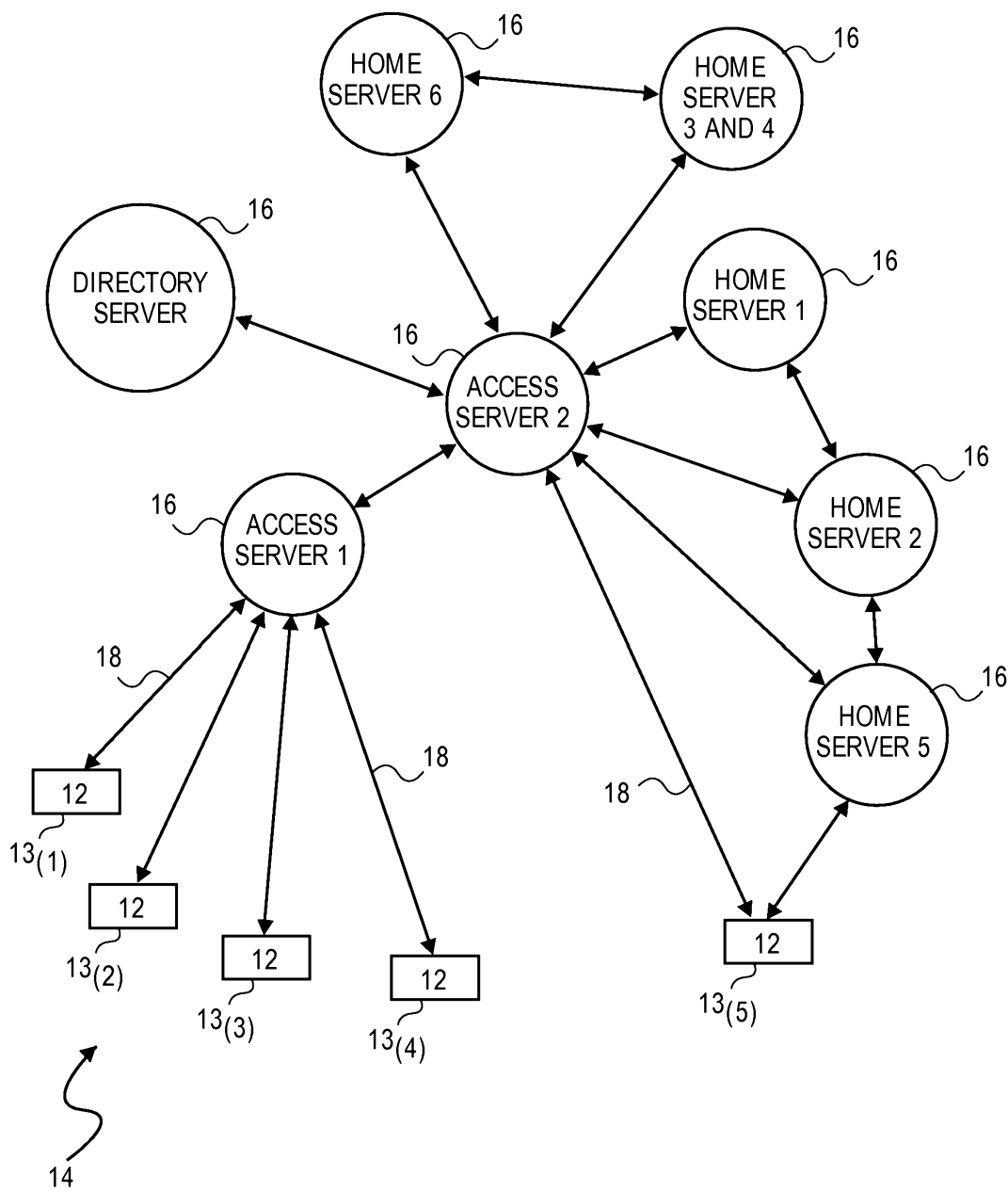
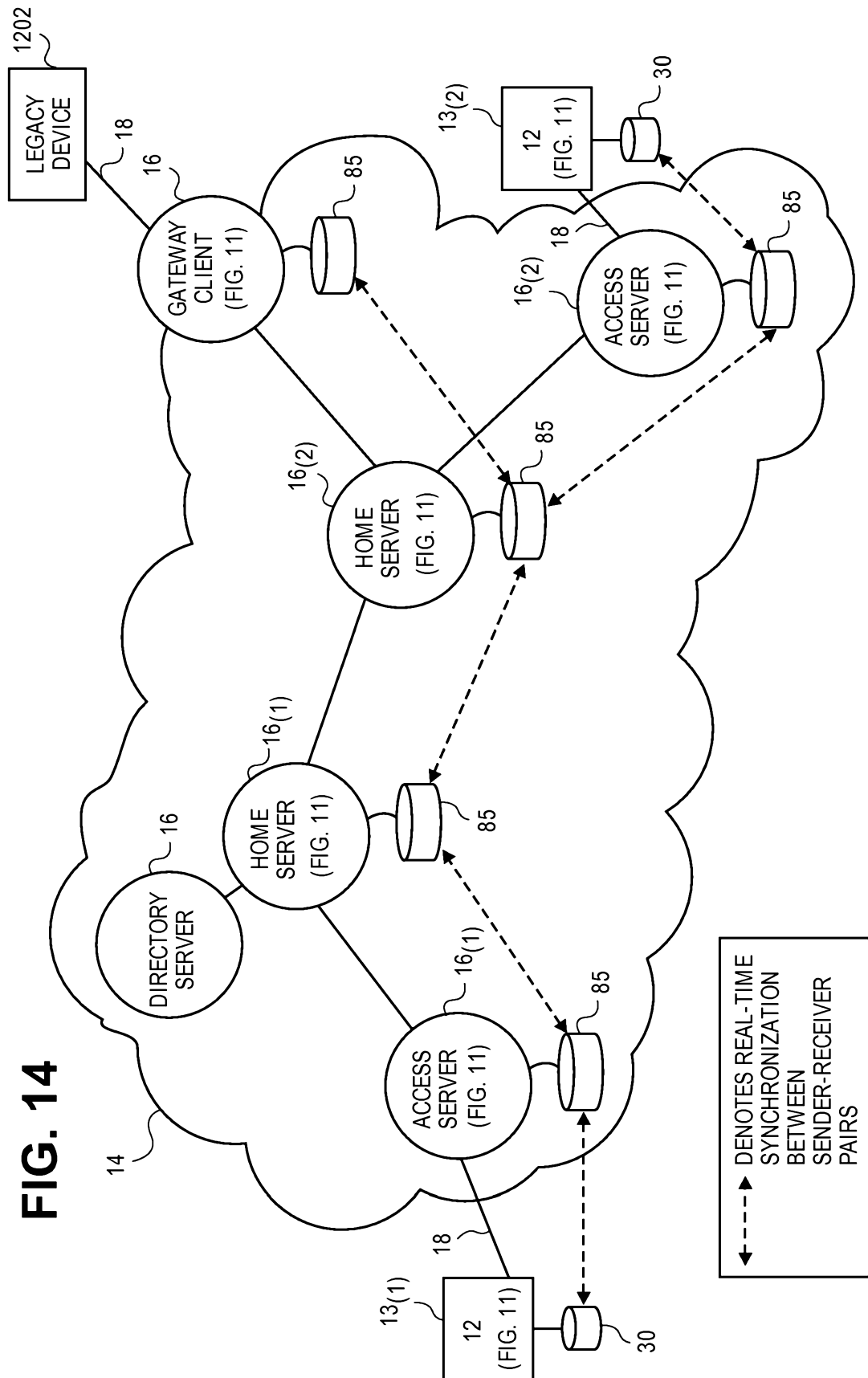


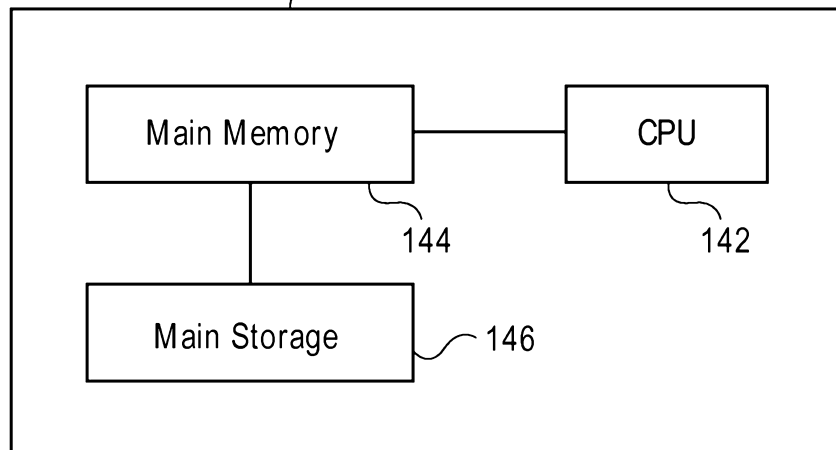
FIG. 13

FIG. 14



26 of 26

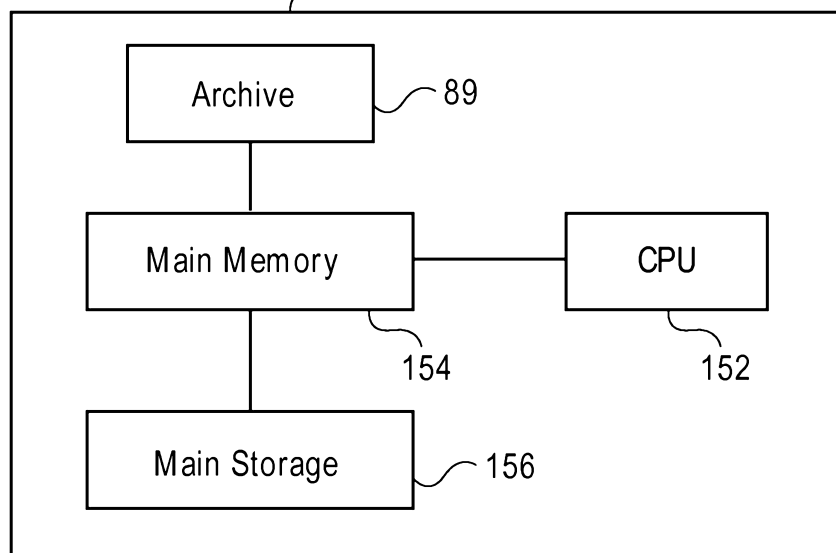
DEVICE 13 Running Client Application 12



140

FIG. 15A

SERVER 16 Running Server Application 78



150

FIG. 15B