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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 March 2016 (17.03.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/040366 A1

(51) International Patent Classification:
H04N 7/14 (2006.01)

(21) International Application Number:
PCT/US2015/049011

(22) International Filing Date:
8 September 2015 (08.09.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
14/480,091 8 September 2014 (08.09.2014) US

(71) Applicant: **EDIFIRE LLC** [US/US]; 245 Summer Street, Boston, MA 02210 (US).

(72) Inventors: **ANDERSON, Eric**; 4 Virginia Circle, Grafton, MA 01519 (US). **GOEPP, Daniel, P.**; 110 Beverly Street, Apt. 1028, Boston, MA 02114 (US).

(74) Agent: **NIEDERMEIER, Patrick, J.**; Proskauer Rose LLP, One International Place, Boston, MA 02110 (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, BN, BR, BW, BY, BZ, CA, CH, CL, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

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

— *with information concerning one or more priority claims considered void (Rule 26bis.2(d))*

(54) Title: METHODS AND SYSTEMS FOR MULTI-FACTOR AUTHENTICATION IN SECURE MEDIA-BASED CONFERENCING

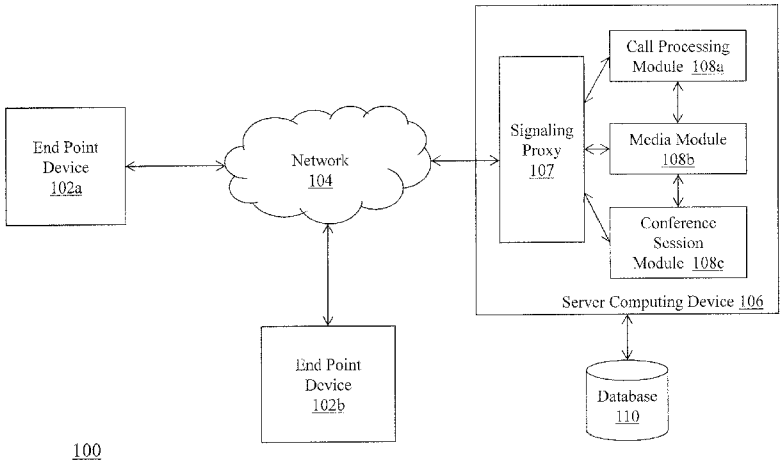


FIG. 1

(57) Abstract: Methods and apparatuses, including computer program products, are described for multi- factor authentication of media-based (e.g., video and/or audio) conferencing between a plurality of end point devices. The methods and apparatuses provide for analysis of an end point media stream using a matrix of authentication factors, where the authentication factors include at least two of: user-specific facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream, to determine an authentication score for the first end point device.

WO 2016/040366 A1

METHODS AND SYSTEMS FOR MULTI-FACTOR AUTHENTICATION IN SECURE MEDIA-BASED CONFERENCING

FIELD OF THE INVENTION

[0001] This application relates generally to methods and apparatuses, including computer program products, for multi-factor authentication in secure media-based conferencing.

BACKGROUND

[0002] Media-based (i.e., video and/or audio) conferencing has traditionally lacked security, call enhancement, and interoperability features. Typically, conference calls occur over private networks between end points that are known to the network and call conferencing system, and have already been authenticated. In some cases, these private networks operate on a proprietary software and/or hardware platform from a specific vendor, making it difficult for end points outside of the scope of the proprietary technology to access the call conferencing platform. In addition, typical conference call signaling is exchanged directly between the end points, which makes security of the signaling harder to achieve.

SUMMARY OF THE INVENTION

[0003] Therefore, what is needed are methods and systems to conduct secure media-based conferencing while offering a robust feature set that provides call enhancement features such as interactive voice response (IVR) functionality and auto attendance, call security features such as password management, multi-factor authentication and authorization of end points (including filtering and allow/deny functionality), and call compliance features such as recording options, regulatory rules, and other retention / surveillance features. The systems and methods described herein provide the advantage of interrupting the signaling associated with a media-based conference call to provide dynamic decision-making with respect to

routing, permissions, authentication, and the like. The systems and methods described herein provide the advantage of secure multi-factor authentication of end points connecting to the call conferencing system. The systems and methods described herein provide the advantage of media-based conferencing with seamless interoperability between end points operating on varying types of software and hardware platforms.

[0004] The systems and methods described herein provide the advantage of dynamically allocating hardware and software resources of the system to ensure availability and efficient routing of media-based conference calls. For example, the hardware and software resources allocated by the systems and methods described herein can be resident on a plurality of geographically-dispersed and independent nodes (*e.g.*, not in the same physical area) that communicate via a mesh-based framework. This attribute enables the system to provide the advantage of a componentized call system, instead of the traditional 'single-box' systems commonly used. Further, the processes and modules that comprise the system can operate independently of each other, without relying on other nodes or modules of the system, in making decisions about which actions to take with respect to a given conference call, end point device, or user. The systems and methods described herein achieve the advantage of disassociating the 'conference call' from any specific physical infrastructure.

[0005] The systems and methods described herein provide the further advantage of enabling private sub-conferences between participants to a main conference call while applying each of the authentication, enhancement, and regulatory features set forth above.

[0006] The invention, in one aspect, features a computerized method for authenticating an end point device participating in a media-based conference call. A call processing module of a server computing device receives a request to join a conference call between a plurality of end point devices. The request includes a media stream associated with the first end point device. The call processing module analyzes the media stream using a matrix of authentication factors, where the authentication factors include at least two of: user-specific

facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream. The call processing module determines an authentication score for the first end point device based upon the media stream analysis and determines whether to connect the first end point device to the conference call, another media resource, or another user of a network or communication system based upon the authentication score.

[0007] The invention, in another aspect, features a system for authenticating an end point device participating in a media-based conference call. The system includes a server computing device having a call processing module configured to receive a request to join a conference call between a plurality of end point devices. The request includes a media stream associated with the first end point device. The call processing module is further configured to analyze the media stream using a matrix of authentication factors, where the authentication factors include at least two of: user-specific facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream. The call processing module is further configured to determine an authentication score for the first end point device based upon the media stream analysis and determine whether to connect the first end point device to the conference call, another media resource, or another user of a network or communication system based upon the authentication score.

[0008] The invention, in another aspect, features a computer program product, tangibly embodied in a non-transitory computer readable storage device, for authenticating an end point device participating in a media-based conference call. The computer program product includes instructions operable to cause a call processing module of a server computing device to receive a request to join a conference call between a plurality of end point devices. The

request includes a media stream associated with the first end point device. The computer program product includes further instructions operable to cause the call processing module to analyze the media stream using a matrix of authentication factors, where the authentication factors include at least two of: user-specific facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream. The computer program product includes further instructions operable to cause the call processing module to determine an authentication score for the first end point device based upon the media stream analysis and determine whether to connect the first end point device to the conference call, another media resource, or another user of a network or communication system based upon the authentication score.

[0009] Any of the above aspects can include one or more of the following features. In some embodiments, the call processing module determines an identity of a user of the first end point device based upon the media stream analysis and retrieves a user profile based upon the identity of the user, the user profile including a set of permissions associated with conference call functionality. In some embodiments, the acoustic environment attributes include (i) one or more pre-recorded audio files from a user and (ii) acoustic attributes of a room in which the first end point device is located. In some embodiments, the visual environment attributes include (i) user-provided or system-captured still images and multiple video frames and (ii) lighting attributes of a room in which the first end point device is located.

[0010] In some embodiments, the technical attributes of the first end point device include an image resolution of a camera coupled to the first end point device, a device identifier, a location, and an originating address. In some embodiments, the technical attributes of the media stream include a media format and a media quality.

[0011] In some embodiments, the call processing module transmits the media stream associated with the first end point device to one or more other end point devices connected to the conference call, receives a validation signal from the one or more other end point devices to confirm an identity of a user of the first end point device, and adds the received validation signal to the matrix of authentication factors. In some embodiments, the media stream analysis includes comparing the matrix of authentication factors for the media stream to a matrix of authentication factors for prior media streams associated with the first end point device. In some embodiments, the media stream analysis includes analyzing the matrix of authentication factors for the media stream based upon user- and/or system-specified preferences.

[0012] In some embodiments, the call processing module stores the media stream analysis and authentication score for the media stream in a database. In some embodiments, the media stream analysis and authentication score determination occur periodically during the conference call. In some embodiments, where the authentication score is below a predetermined threshold, the call processing module transmits a request for user credentials to the first end point device. In some embodiments, the call processing module transmits the authentication score for the first end point device for display on one or more other end point devices connected to the conference call.

[0013] In some embodiments, the call processing module denies the request to join the conference call when the authentication score is below a predetermined threshold and disconnects the first end point device. In some embodiments, the call processing module allows the request to join the conference call when the authentication score is at or above a predetermined threshold and connects the first end point device to the conference call.

[0014] Other aspects and advantages of the invention will become apparent from the following detailed description, taken in conjunction with the accompanying drawings, illustrating the principles of the invention by way of example only.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] The advantages of the invention described above, together with further advantages, may be better understood by referring to the following description taken in conjunction with the accompanying drawings. The drawings are not necessarily to scale, emphasis instead generally being placed upon illustrating the principles of the invention.

[0016] FIG. 1 is a block diagram of a system for media-based conferencing between a plurality of end point devices, according to an embodiment of the invention.

[0017] FIG. 2 is a flow diagram of a method for media-based conferencing between a plurality of end point devices, according to an embodiment of the invention.

[0018] FIG. 3 is a block diagram of a system for media-based conferencing between a plurality of end points, according to an embodiment of the invention

[0019] FIG. 4 is a flow diagram of a method for authenticating an end point device participating in a conference call, according to an embodiment of the invention.

[0020] FIG. 5 is a block diagram of a system for establishing a sub-conference between a plurality of end point devices participating in a conference call, according to an embodiment of the invention.

[0021] FIG. 6 is a flow diagram of a method for establishing a sub-conference between a plurality of end point devices participating in a conference call, according to an embodiment of the invention.

[0022] FIG. 7 is a block diagram of a networked system for media-based conferencing between a plurality of end points, according to an embodiment of the invention.

[0023] FIG. 8 is a flow diagram of a multi-factor authentication process of an end point device participating in a conference call, according to an embodiment of the invention.

DETAILED DESCRIPTION

[0024] FIG. 1 is a block diagram of a system 100 for media-based conferencing between a plurality of end point devices, according to an embodiment of the invention. The system 100 includes a plurality of end point devices 102a-102b, a communications network 104, a server computing device 106 that includes a signaling proxy 107, a call processing module 108a, a media module 108b, and a conference session module 108c, and a database 110.

[0025] The plurality of end point devices 102a-102b connect to the server computing device 106 via the communications network 104 in order to initiate and participate in conference calls and other media communication sessions with other end points. Exemplary end point devices include desktop computers, laptop computers, tablets, mobile devices, smartphones, and internet appliances. It should be appreciated that other types of computing devices that are capable of connecting to the server computing device 106 can be used without departing from the scope of invention. In some embodiments, the end point devices 102a-102b are capable of executing call conferencing client software locally and/or using another type of user interface (*e.g.*, a web browser) to connect to the server computing device 106. The call conferencing client software can be open network, free-to-use / freemium software, such as Skype™ available from Microsoft Corp. of Redmond, Washington or Google™ Hangouts available from Google, Inc. of Mountain View, California, or purchasable, closed network software, such as the RealPresence® platform available from Polycom, Inc. of San Jose, California. In some embodiments, the call conferencing client software can be a proprietary platform developed, *e.g.,* by a corporation for use internally. Although FIG. 1 depicts two end point devices 102a-102b, it should be appreciated that the system 100 can include any number of end point devices.

[0026] The communication network 104 enables the end point devices 102a-102b to communicate with the server computing device 106 in order to initiate and participate in media-based conference calls and meetings. The network 104 may be a local network, such

as a LAN, or a wide area network, such as the Internet and/or a cellular network. In some embodiments, the network 104 is comprised of several discrete networks and/or sub-networks (*e.g.*, cellular → Internet) that enable the end point devices 102a-102b to communicate with the server computing device 106.

[0027] The server computing device 106 is a combination of hardware and software modules that establish, authorize, facilitate and manage media-based conference calls and meetings between a plurality of end point devices 102a-102b. The server computing device 106 includes a signaling proxy 107, a call processing module 108a, a media module 108b, and a conference session module 108c. The proxy 107 and modules 108a-108c are hardware and/or software modules that reside on the server computing device 106 to perform functions associated with establishing, authorizing, facilitating, and managing media-based conference calls and meetings. In some embodiments, the functionality of the proxy 107 and the modules 108a-108c is distributed among a plurality of computing devices. It should be appreciated that any number of computing devices, arranged in a variety of architectures, resources, and configurations (*e.g.*, cluster computing, virtual computing, cloud computing) can be used without departing from the scope of the invention. It should also be appreciated that, in some embodiments, the functionality of the proxy 107 and the modules 108a-108c can be distributed such that any of the proxy 107 and/or the modules 108a-108c are capable of performing any of the functions described herein without departing from the scope of the invention. For example, in some embodiments, the functionality of the proxy 107 and/or the modules 108a-108c can be merged into a single module or, in some embodiments, the modules 108a-108c can be merged into a single module coupled to the proxy 107.

[0028] The signaling proxy 107 is coupled between the end point devices 102a-102b and the modules 108a-108c. The proxy 107 receives signaling communications in various protocols (*e.g.*, Session Initiation Protocol (SIP), h.323) from the end point devices 102a-102b that relate to the establishment and maintenance of media-based conference call sessions. It

should be appreciated that other signaling protocols can be used without departing from the scope of invention. The proxy 107 receives the signaling communications from the end point devices 102a-102b and transmits the signaling to the modules 108a-108c for further processing.

[0029] In some embodiments, the proxy 107 translates the signaling into another format that can be processed by the modules 108a-108c. For example, the proxy 107 can translate the signaling into XML format, including items such as call details (*e.g.*, to, from, domain), end point device-specific details, user-specific details, and other types of information, thereby offering additional customization of the signaling to enable the modules 108a-108c to process the call in a dynamic way for each end point device 102a-102b.

[0030] Upon receiving the signaling from the proxy 107, the call processing module 108a can perform a number of different actions to process the call. In some embodiments, the call processing module 108a analyzes the signaling and redirects the call to other resources in the system 100 for further processing. For example, the call processing module 108a can determine that the inbound call signaling is originating from an end point device that is operating a specific call conferencing hardware and/or software platform. Based upon the platform determination, the call processing module 108a can redirect the signaling to a resource in the system that is capable of communicating with the platform of the end point device 102a. In some embodiments, the call processing module 108a returns a response to the end point device 102a that originated the signaling, where the response includes call routing data (*e.g.*, a URI) for the end point device to re-route the signaling. In some embodiments, the call processing module 108 returns a response to the proxy 107 (*e.g.*, XML) and the proxy 107 handles the routing of the signaling session to the destination end point device.

[0031] In some embodiments, the call processing module 108a uses the signaling to identify a user of the originating end point device 102a and/or the type of end point device 102a that originated the signaling. For example, the call processing module 108a can utilize

data in the signaling, such as the 'to' address, the 'from' address, a device identifier, a user ID, and the like, to determine the identity of a user associated with the originating end point device or the destination end point device. The call processing module 108a can access the database 110 to look up details of the user based upon any of the above data points. For example, if the signaling includes a 'to' address, the call processing module 108a can search in the database 110 for a user profile associated with the 'to' address. In this way, the call processing module 108a maps the signaling to a user and can then leverage its capabilities to customize the conference experience based upon that user's identity.

[0032] In another example, the call processing module 108a can use the signaling to determine the technical capability of the end point device 102a and adjust the conferencing features and options available to that end point device. The signaling can include a data point that indicates the originating end point device 102a has limited network bandwidth for sending and receiving data. The call processing module 108a can upgrade or downgrade the fidelity of the media transmitted to the originating end point device 102a based upon the available bandwidth capabilities of the device 102a.

[0033] In another example, the call processing module 108a can use the signaling to determine a user associated with the end point device (as described above) and then perform authentication of the end point device / user to determine the level of access that the user has on the system 100. For example, the call processing module 108a can determine that the user is restricted from establishing media-based conference calls with a specified list of destinations (*e.g.*, people, devices, physical locations). Based upon the determination of these restrictions, the call processing module 108a can evaluate whether to establish the conference call between the originating end point device 102a and the destination end point device specified in the signaling.

[0034] As described above, the server computing device 106 also includes a media module 108b. The media module is coupled to the proxy 107 and the other modules 108a and

108c. The media module 108b performs media signaling and streaming functionality, including acting as a call termination and streaming point. In some embodiments, the media module 108b performs functions to locate an intermediate point (*e.g.*, server device) between the participating end point devices to anchor the media and may not process the media flow. In some embodiments, once the call processing module 108a has established the conference call between a plurality of end point devices based upon the signaling, the call processing module 108a can transfer the media associated with the call to the media module 108b for handling the media session between the end point devices. The media module 108b also provides additional conference call enhancement features, such as Interactive Voice Response (IVR) menus and prompts, auto-attendants, and advanced PIN management.

[0035] In some embodiments, the media module 108b includes private branch exchange (PBX) software for managing the media signaling and flows of conference calls processed by the server computing device 106. An example PBX software platform that can be incorporated into the media module is Asterisk®, available from Digium, Inc. of Huntsville, Alabama.

[0036] The server computing device 106 also includes a conference session module 108c. The conference session module 108c is coupled to the proxy 107 and the other modules 108a and 108b. The conference session module 108c performs functions to bridge the plurality of end point devices participating in a media-based conference call or meeting into a single session. In some embodiments, the conference session module 108c is a multipoint control unit (MCU). An example MCU that can be incorporated into the conference session module 108c is Codian, available from Cisco Systems, Inc. in San Jose, California. The MCU can be integrated with Vidtel Gateway, available from Vidtel, Inc. of Sunnyvale, California, to provide additional features.

[0037] The system 100 also includes a database 110. The database 110 is coupled to the server computing device 106 and stores data used by the server computing device 106 to

perform the media-based conferencing functionality. The database 110 can be integrated with the server computing device 106 or be located on a separate computing device. An example database that can be used with the system 100 is MySQL™ available from Oracle Corp. of Redwood City, California.

[0038] FIG. 2 is a flow diagram of a method 200 for media-based conferencing between a plurality of end point devices, using the system 100 of FIG. 1. The call processing module 108a of the server computing device 106 receives (202) a request to establish a conference call between a plurality of end point devices (*e.g.*, 102a-102b). The request originates from one of the end point devices (*e.g.*, 102a). For example, in the SIP context, the originating end point device 102a transmits a SIP INVITE message to the server computing device 106 via the network 104. The proxy 107 receives the SIP INVITE message and, in some embodiments, translates the SIP INVITE message into XML, which is then passed on to the call processing module 108a.

[0039] The call processing module 108a allocates (204) a conference call resource identifier to the originating end point device 102a, based upon the received XML request. As will be described in greater detail below, the call processing module 108a also determines the availability of resources for the conference call and transmits the conference call resource identifier to the originating end point device 102a only if specified and/or necessary resources are available.

[0040] The call processing module 108a determines (206) conference call attributes based upon one or more of: technical specifications of the originating end point device 102a, an identity of a user associated with the originating end point device 102a, a geographic location of the originating end point device, and the request to establish the conference call. As described above, the call processing module 108a receives the request in XML format from the proxy 107 and evaluates data embedded in the request to determine processing for the conference call request. For example, the call processing module 108a can evaluate the

origination address (*e.g.*, 'from' address) in the request to identify a user of the end point device 102a in order to perform functions such as authentication, group permissions, feature access, and the like. In another example, the call processing module 108a can evaluate the request to determine certain technical attributes of the originating and/or destination end point devices. For instance, if the call request originates at a mobile device and/or via a cellular network, the call processing module 108a can determine extra and/or different call security options (*e.g.*, encryption). In another example, the call processing module 108a can evaluate the geographic location of the originating end point device 102a to determine appropriate routing rules, security authorizations, and/or hardware / software resources to allocate to the originating end point device 102a or the conference call itself. It should be appreciated that the call processing module 108a can analyze the request to establish a conference call to determine attributes associated with the call and/or attributes that the module 108a should assign to the call as the request arrives.

[0041] As mentioned above, the call processing module 108a determines (208) an availability of resources based upon the conference call attributes. For example, the call processing module 108a determines a resource (*e.g.*, gateway, URI) available for the requested conference call and generates a conference call resource identifier to be assigned to the originating end point device 102a. For example, if the originating end point device 102a is using a particular software platform (*e.g.*, Skype™) to initiate the media-based conference call, the call processing module 108a generates a Skype™ URI to be associated with the end point device 102a. In some embodiments, if a particular resource is unavailable (*e.g.*, all of the Skype™ URIs are in use), the call processing module 108a can inform the originating end point device 102a and wait until a resource is available or decline to establish the conference call.

[0042] The call processing module 108a can then transmit (210) the conference call identifier (*e.g.*, URI) to the end point device 102a for redirection of the call signaling to a

server associated with the software platform. In some embodiments, the conference call resource identifier is associated with resources internal to the system 100 that can handle and process the call signaling (*e.g.*, an internal gateway, transcoder).

[0043] The call processing module 108a transmits (212) the determined conference call attributes to the media module 108b. For example, the call processing module 108a can transmit the determined conference call attributes (*e.g.*, via XML) to the media module 108b for providing features such as call enhancement functionality, call routing functionality, media streaming functionality, and call termination functionality.

[0044] The media module 108b establishes (214) a conference media connection between the originating end point device 102a and the conference session module 108c for initiation of a conference media flow. In some embodiments, once the call processing module 108a has established the signaling connection and performed various functions (*e.g.*, authentication, permissioning, routing), the media module 108b initiates a media flow between the originating end point device 102a and the conference session module 108c at the server computing device 106. The media module 108b can still manage the call state and media streaming, while the conference session module 108c connects (216) the various end point devices participating in the conference call to the conference call session. In some embodiments, the conference session module 108c uses the determined conference call attributes to determine whether certain end point devices are able to connect to the conference call.

[0045] FIG. 3 is a block diagram of a system 300 for media-based conferencing between a plurality of end points, based upon the system 100 of FIG. 1. The system 300 includes the end point device 102a, the server computing device 106, the database 110, a third-party call conferencing server computing device 302, a call conferencing transcoder 304, a media-based communication server (MCS) computing device 306, and a MCU 308.

[0046] FIG. 3 depicts an example workflow for establishing a media-based conference call between a plurality of end point devices:

[0047] Step 1: the end point device 102a transmits a request to establish a media-based conference call to the server computing device 106. The request includes a meeting ID. For example, if the end point device 102a is operating via Skype™ client software, a user at the device 102a clicks a button in the Skype™ user interface to initiate the conference call. The software transmits the user's Skype™ ID and the meeting ID to the server computing device 106 to establish a signaling session with the server 106.

[0048] Step 2: the server computing device 106 processes the request to determine the identity of the user associated with the request. For example, the server 106 can retrieve a user ID from the request and use the database 110 to determine the identity of the user and any related information (*e.g.*, permissions, device-specific information, and authentication details). The server can check whether the user associated with the user ID is authorized to join the meeting requested (via the meeting ID). The server computing device 106 then returns a Skype™ URI to the end point device 102a.

[0049] Step 3: the end point device 102a uses the received URI to initiate a signaling session for the conference call with a Skype™ server (*e.g.*, third-party call conferencing server 302). The URI can include an address or identifier associated with the call conferencing transcoder (*e.g.*, device 304), the media-based communication server (*e.g.*, device 306), and/or the MCU 308. In some embodiments, the URI corresponds to an MCS endpoint and the server computing device 106 / database 110 maintains a correspondence between the third-party URI and the MCS endpoint.

[0050] Step 4: the third-party call conferencing server 302 forwards the conference call signaling to the call conferencing transcoder 304 (*e.g.*, a Vidtel module).

[0051] Step 5: the call conferencing transcoder 304 maps the received Skype™ URI to an MCS endpoint address. For example, the transcoder 304 can modified the user's Skype™ ID

to add the MCS end point address (*e.g.*, <user's Skype™ ID>@skype.vidtel.com). The transcoder 304 then communicates with the MCS 306.

[0052] Step 6: the MCS 306 communicates with the server computing device 106. For example, the MCS 306 transmits the modified Skype™ ID that includes the MCS end point address to the server computing device 106.

[0053] Step 7: the server 106 uses the modified ID to locate the meeting ID that was previously transmitted to the server 106 by the end point device 102 (see Step 1). The server 106 then transmits the meeting ID to the MCS 306.

[0054] Step 8: the MCS 306 then transfers the conference call signaling for the end point device 102a to the MCU 308 so that the user can join the media-based conference call or meeting, as requested.

[0055] It should be appreciated that FIG. 3 represents an exemplary embodiment for establishing a media-based conference call between a plurality of end point devices. Other techniques and workflows for establishing a media-based conference call between a plurality of end point devices can be contemplated without departing from the scope of invention.

[0056] In some embodiments, the system 100 is capable of providing interoperability between different types of end point devices that request to join the same media-based conference call. For example, an end point device 102a may access the system 100 via a Skype™ user interface over a cellular network while another end point device 102b may access the system via a call conferencing terminal installed behind a VoIP architecture, *e.g.*, at a company. The server computing device 106 can perform transcoding between the respective end point devices to enable a seamless communication session. In some embodiments, the server computing device 106 reserves system resources depending on the type of request and/or end point device that is accessing the system. The server computing device 106 can maintain a pool of access so that system resources are reserved and released as

appropriate based upon the initiation and termination of media-based conference calls and meetings.

[0057] FIG. 4 is a flow diagram of a method for authenticating an end point device participating in a media-based conference call, using the system 100 of FIG. 1. The call processing module 108a of the server computing device 106 receives (402) receives a request to establish a conference call between a plurality of end point devices (*e.g.*, end point devices 102a-102b), the request including credentials associated with a user of an end point device (*e.g.*, 102a) and attributes associated with the end point device 102a. The call processing module 108a determines (404) an identity of the user of the end point device based upon the credentials. The call processing module 108a determines (406) a level of conference call access based upon the attributes associated with the end point device. The call processing module 108a retrieves (408) a user profile based upon the identity of the user, the user profile including a set of permissions associated with conference call functionality. The call processing module 108a determines (410) whether to connect the end point device to a conference call based upon the user profile and the level of conference call access.

[0058] For example, the server computing device 106 can perform multi-factor authentication to determine (i) the identity of both the user and end point device that has initiated the request to establish a conference call and (ii) how the server 106 should handle the request (*e.g.*, connect the call, route the call to another resource in the system for further processing, authenticate the user via additional means). In one example, the server computing device 106 can fingerprint the incoming request by evaluating attributes relating to a user of the end point device, the end point device itself, and/or the parameters of the requested call to create a matrix of the attributes that comprise the fingerprint. The server 106 can store the fingerprint, *e.g.*, in the database 110 so that the fingerprint can be referenced in the future – both in comparison with requests from the same user / end point device and from different users / end point devices. The server computing device 106 can also re-evaluate the

fingerprint as a conference call proceeds after establishment. For example, if the end point device is attempting to access the conference call system in an unauthorized manner by providing a temporary, fake alias to the server 106 upon transmitting the call request, but later during the call the end point device attributes change, the server 106 can dynamically and automatically evaluate the fingerprint again and take corrective action (*e.g.*, disconnecting the call, routing the call to another resource for monitoring) if necessary.

[0059] In some embodiments, the multi-factor authentication is implemented according to two facets: (1) technically (*i.e.*, how/why is the end point device calling?) and (2) personally (*i.e.*, who is the user at the end point device?). The server computing device 106 can receive a request for a media-based conference from an end point device and determine items such as technical features of the calling device (*e.g.*, hardware, software, location, originating network, protocol) and personal features of the calling device and/or the user associated with the device (*e.g.*, user name, ID, PIN, facial recognition). The server computing device 106 can evaluate both sets of information to provide a robust authentication process that ensures appropriate access and permissions are granted to the end point device.

[0060] The system 100 can perform multi-factor authentication based upon a matrix of user-specific authentication factors, including but not limited to: facial recognition attributes; audio recognition (*e.g.*, speech patterns / wavelengths) attributes; acoustic environment attributes; visual environment attributes; user gesture attributes; technical attributes of the end point device being operated by a user; and technical attributes of the media stream received from the end point device. Exemplary acoustic environment attributes include acoustics of the room / environment in which the user is located during the call and acoustic attributes of the equipment (*e.g.*, microphone) or device used to capture audio during the call. Exemplary visual environment attributes include objects in the environment around the user (*e.g.*, walls, windows, paintings, furniture), colors of the environment, lighting hues of the environment, and the like.

[0061] Exemplary user gesture attributes include placement and/or movement of the user's body during the call (*e.g.*, the user's hand movement while speaking), a predetermined gesture pattern performed by the user at the beginning of the call, and other similar types of gesture recognition attributes. Exemplary technical attributes of the end point device being operated by a user include the hardware and/or software platform from which the user is connecting to the call, attributes of the image-capturing device, and address / identification attributes of a device being operated by the user (*e.g.*, IP address, MAC address, geolocation / GPS characteristics, and the like). Exemplary technical attributes of the media stream received from the end point device include bit rate, jitter, delay, compression protocol / media format / codec utilized, transport protocol, encryption standard, and the like.

[0062] The system 100 can analyze any of the above-described authentication factors to determine an authentication score for the end point device connecting to the call. The authentication score represents a level of confidence that the end point device and/or user requesting to join a call is indeed authorized to do so. In some embodiments, the authentication score is determined by comparing the current matrix of authentication factors against historical authentication factors associated with the end point device (*e.g.*, stored in a database) in order for the system to determine whether or not to connect the end point device to the call. In some cases, the system 100 can generate a threshold value based upon the historical authentication factors. For example, if the authentication score falls below the threshold, the system 100 can decide that the end point device is unauthorized and prevent the device from connecting to the call (and/or disconnect the device from the system 100).

[0063] In another example, if the authentication score falls below the threshold, the system 100 can allow the end point device to connect to the call and also alert other users on the call that the end point device / user is unauthenticated. The other users may be able to authenticate the end point device / user by validating the identity of the user (*e.g.*, using a validation signal) and informing the system 100 (i) that the user is authorized to participate

and/or (ii) instructing the system 100 of the user's identity. In the latter case, the system 100 is capable of storing the current authentication factors in the user's profile (*e.g.*, as a secondary set of authentication factors) – such that if the user / end point device requests to join a future call, the system 100 can recognize the user / end point device based upon the secondary set of authentication factors as part of a self-learning process. In some cases, the system 100 can merge the secondary set of authentication factors with authentication factors already stored for the user / end point device to result in an updated user profile.

[0064] The system 100 can display a user's / end point device's authentication score on the call screen to the other users so everyone is aware of the others' scores. In one embodiment, the score is displayed as a percent-based number (*e.g.*, 75%) and can be displayed using color coding. For example, authentication scores above 75% are colored green, authentication scores between 50 and 75% are colored orange, and authentication scores below 50% are colored red.

[0065] In some embodiments, the above-described multi-factor authentication techniques are performed at the initiation of a call (or a request to join a call) and periodically during the call. The authentication that occurs during the call can be pre-scheduled (*e.g.*, every fifteen minutes) or conducted dynamically based upon changes to the matrix of authentication factors. For example, if an end point device participating in a call changes its IP address / location (*e.g.*, due to a cell phone moving geographically), the system 100 can flag that end point device for a real-time authentication. In another example, a user may pull a shade down in the room to reduce glare, thereby changing the lighting hues in the user's video. The system 100 can perform a real-time authentication of the end point device to determine whether the device is still authorized to participate in the call.

[0066] FIG. 8 is a flow diagram of a multi-factor authentication process of an end point device participating in a conference call. It should be appreciated that any of the process

functions 802, 804, 806, 808, 810_described in FIG. 8 can be performed by one or more of the modules 107, 108a-108c described in the system 100 of FIG. 1.

[0067] An end point device 102a operated by a user transmits a request to the server computing device 106 to join an in-progress video conference call. The call from end point device 102a is received by the server computing device 106 at the core call control function 802. The core call control function 802 communicates with the decision engine 804 to determine that the video conference which the end point device wants to join is a call that requires multi-factor authentication. The core call control function 802 communicates with the interactive voice response (IVR) function 806 which asks the user at the end point device 102a to wait while authentication is performed. The matrix of authentication factors about the end point device 102a (as described previously) is transmitted to the multi-factor authentication service 808 which performs the analysis described above to generate an authentication score for the end point device 102a. If the multi-factor authentication service 808 determines that the end point device 102a is authorized to join the call, the core call control function 802 transfers the media stream of the end point device 102a to a multi-party conferencing unit (*i.e.*, MCU 810) on which the video conference call is taking place. In some embodiments, the core call control function 802 can transmit the end point device 102a to a waiting room (*e.g.*, if further authentication is required, such as a passcode, or if the organizer has not yet joined). When the end point device 102a joins the video conference call, the MCU 810 can transmit data relating to the end point device 102a and its media stream to the multi-factor authentication service 808 in order to conduct the periodic / dynamic authentication procedure and the self-learning procedure described earlier. This data can be stored in a database 110 for future use.

[0068] The system 100 is also capable of performing authentication both at an individual permission level and a call / meeting permission level. For example, the server computing device 106 can retrieve individual permissions from the database 110, *e.g.*, based upon a

Lightweight Directory Access Protocol (LDAP) or Active Directory (AD) service. The server computing device 106 can tie in to an organization's internal user profile and directory service to use the user information for managing call conferencing permissions. The server computing device 106 can also maintain and manage call / meeting permissions, such as generating or storing a list of participants that are eligible to join a particular conference call or meeting, *e.g.*, based upon user identifier, job title, access level, or other such parameters as may be defined in relation to the user.

[0069] The system 100 has a robust permissioning scheme that allows for customized authentication and permissions at a functional level, a personal level, a group level, and/or a device level. For the functional permissions, the system 100 can determine that a particular user / end point device / conference call is able to execute certain functions, *e.g.*, set up a sub-conference, establish a conference call, invite participants, implement security attributes, and so forth. For the personal permissions, the system 100 enables a granularity of access controls for a particular user. For example, the system 100 can determine that a user has permission to initiate a conference call of a particular type, from a particular location (*e.g.*, work but not home), to particular people, and the like.

[0070] For the group permissions, the system 100 can assign attributes to a plurality of users based upon any number of different requirements or categorizations. For example, all of the employees in a corporation can be assigned to a group. The system 100 can also assign users to multiple groups and/or sub-groups, such as business units within the corporation or product teams within the business unit. Users can also form their own ad-hoc groups (*e.g.*, friends, business contacts) comprising any number of users, *e.g.*, from two users and up. It should be noted that the users do not have to be otherwise affiliated (*e.g.*, by employer) in order to be in a group. For the device permissions, the system 100 can determine a level of access and functionality for particular call environments (*e.g.*, hardware and/or software

platform (or other technical attributes) of an end point device, location of an end point device, and so forth).

[0071] Another aspect of the permissioning structure described herein is the hierarchical nature of the permissions. For example, a corporation may assign all of its employees to a particular group with specifically-designated permission levels. Certain employees within that group may form a sub-group and want a permission level that differs from the corporation-wide permission structure. The system 100 can determine that the corporation-wide permission structure overrides the sub-group permission level, and enforce limits on the sub-group permission level so as to not exceed the scope of or conflict with the corporate-wide permissions structure.

[0072] In some scenarios, participants to a main conference call may wish to separate themselves from the conference for a brief period to communicate regarding private matters. For example, if the main conference call involves business negotiations or legal proceedings with multiple parties, one party may want to confer regarding sensitive or privileged issues without disconnecting from the main conference but still engaging in a private conversation. Therefore, the system offers a function to establish a sub-conference between multiple end point devices, and the system can apply all of the functionality described above (*e.g.*, authentication, routing, and permissioning) to the sub-conference.

[0073] FIG. 5 is a block diagram of a system 500 for establishing a sub-conference between a plurality of end point devices (*e.g.*, end point devices 102a, 102b, 602a, 602b) participating in a media-based conference call, based upon the system of FIG. 1. The system 500 includes end point devices 102a, 102b, 502a, 502b, and server computing device 106. The end point devices 102a, 102b, 502a, 502b connect to the main conference call 504 using the techniques described above. Once the main conference call 504 is established, certain participants may wish to separate themselves from the main call 504 for a period of time.

[0074] FIG. 6 is a flow diagram of a method 600 for establishing a sub-conference between a plurality of end point devices participating in a media-based conference call, using the system 100 of FIG. 1 and the system 500 of FIG. 5. The server computing device 106 receives (602) a request to establish a sub-conference from a first end point device (*e.g.*, end point device 102a) participating in a main conference call. The request includes an identifier associated with a second end point device (*e.g.*, end point device 102b) participating in the main conference call. For example, the user at end point device 102a may click on an icon associated with a user at end point device 102b to initiate a sub-conference with the user at end point device 102b, and the device 102a transmits a request to the server computing device 106.

[0075] The server computing device 106 initiates (604) a call instance (*e.g.*, separate call instance 506) that is separate from the main conference call 504 based upon the request to establish the sub-conference. For example, the server computing device 106 can initiate separate call instance 506 by allocating another meeting that is managed by the server 106 but where the media flow is separate from the main conference call 504. In some embodiments, only a portion of the media flow is transferred to the separate call instance 506. For example, the participants in the main conference call may continue to see video associated with the users involved in the sub-conference, but the audio communication exchanged between the users involved in the sub-conference is excluded from the main conference call 504. In another example, the participants in the main conference call may hear a tone indicating that some users have initiated a sub-conference. In some embodiments, the sub-conference include the exchange of textual information (*e.g.*, chat) between the participants in the sub-conference.

[0076] In some embodiments, the server computing device 106 changes presence information associated with the users and/or end point devices that joined the sub-conference. For example, the server computing device 106 can add a status indicator to a user's name or

picture in a video conference indicating to all participants in the main conference call that the user is in a sub-conference. In another example, the server computing device 106 can periodically play an audio message to participants in the main conference call that certain users have entered a sub-conference and identify those users by name.

[0077] The server computing device 106 couples (606) the first end point device 102a and the second end point device 102b to the separate call instance 506 without separating the first end point device 102a and the second end point device 102b from the main conference call 504. The server computing device 106 establishes (608) a media flow between the first end point device 102a and the second end point device 102b, where the main conference call 504 is prevented from accessing the media flow between the first end point device 102a and the second end point device 102b.

[0078] The server computing device 106 can apply any of the concepts described above with respect to call routing, enhancement features, permissions, authentication, and the like to the sub-conference call instance 506. For example, if a participant in the main conference call seeks to establish a sub-conference with another participant, the server 106 can determine whether the first participant is allowed to communicate privately with the second participant, *e.g.*, via a set of rules stored in the database 110. The rules can be pre-determined in advance of the main conference call or agreed upon by the participants at the time of the call. The rules can change during the course of the call, *e.g.*, as participants come and go from the main conference call. The permissions and related features can be hierarchical across the main conference call and the sub-conference, so that a user has the same permission scope in each call instance, or in some embodiments, a user may have reduced permission scope in a sub-conference versus his permission scope in a main conference call.

[0079] It should be noted that, in some embodiments, the server computing device 106 keeps track of the status of each participant in the conference call. For example, the server 106 records certain data about the circumstances in which the users entered the sub-

conference (*e.g.*, time at which the sub-conference was initiated, identify of users in the sub-conference, how long the sub-conference lasted, and the like). This feature allows for detailed reporting about the structure and timing of the conference call, which is useful for audit and compliance purposes.

[0080] In certain contexts, the rules provide for certain surveillance and reporting features to be applied to the sub-conferences. For example, an industry may be regulated such that all communications must be recorded and documented according to specific guidelines. The server computing device 106 can determine, based upon the rules, whether to record the sub-conferences, identify the participants to the sub-conferences, and other similar requirements. In some cases, the rules applied to the main conference call are automatically transferred to the sub-conferences. In addition, the system can provide additional features such as surveillance or reporting to be used in conjunction with the sub-conference.

[0081] In addition, an advantage provided by the methods and systems described herein is that each end point device (*e.g.*, device 102a of FIG. 1) has its own media stream when participating in a conference call. Accordingly, the server computing device 106 and/or other intermediary servers and devices can insert artifacts that are personalized or specific to the user of a certain end point device (*e.g.*, notifications, alerts, messages, graphics) without inserting the artifacts into the media streams of other end point devices. This technique enables the system 100 to perform dynamic and personalized user interface, recording, and display operations for individual users.

[0082] The techniques may be implemented in a networked system 700 comprising multiple computing devices distributed across different locations, as shown in FIG. 7. Each of Location A 702, Location B 704 and Location C 706 includes the server computing device 106 having components 107, 108a-108c, and 110 of FIG. 1, and the servers at locations 702, 704, and 706 are connected to each other via the network 104. The networked system of FIG. 7 enables distribution of the processing functions described herein across several computing

devices and provides redundancy in the event that a computing device at one location is offline or inoperable. In some embodiments, end point devices (*e.g.*, device 102a) in proximity to a particular location (*e.g.*, Location A 702) access the networked system via the server 106 at that location. In some embodiments, the server computing devices 106 at the respective locations 702, 704, 706 communicate with a central computing device 712 (*e.g.*, a server) that is coupled to the network. The central computing device 712 can provide data and/or processing resources for the network of computing devices 106 (*e.g.*, synchronization of functionality/data across the computing devices).

[0083] It should be understood that any of the above-described methods, systems, and techniques can be implemented in the context of video conferencing (*i.e.*, conference calls consisting of video and audio media) and audio-only conferencing without departing from the scope of invention.

[0084] The above-described techniques can be implemented in digital and/or analog electronic circuitry, or in computer hardware, firmware, software, or in combinations of them. The implementation can be as a computer program product, i.e., a computer program tangibly embodied in a machine-readable storage device, for execution by, or to control the operation of, a data processing apparatus, *e.g.*, a programmable processor, a computer, and/or multiple computers. A computer program can be written in any form of computer or programming language, including source code, compiled code, interpreted code and/or machine code, and the computer program can be deployed in any form, including as a stand-alone program or as a subroutine, element, or other unit suitable for use in a computing environment. A computer program can be deployed to be executed on one computer or on multiple computers at one or more sites.

[0085] Method steps can be performed by one or more processors executing a computer program to perform functions of the invention by operating on input data and/or generating output data. Method steps can also be performed by, and an apparatus can be implemented as,

special purpose logic circuitry, *e.g.*, a FPGA (field programmable gate array), a FPAA (field-programmable analog array), a CPLD (complex programmable logic device), a PSoC (Programmable System-on-Chip), ASIP (application-specific instruction-set processor), or an ASIC (application-specific integrated circuit), or the like. Subroutines can refer to portions of the stored computer program and/or the processor, and/or the special circuitry that implement one or more functions.

[0086] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital or analog computer. Generally, a processor receives instructions and data from a read-only memory or a random access memory or both. The essential elements of a computer are a processor for executing instructions and one or more memory devices for storing instructions and/or data. Memory devices, such as a cache, can be used to temporarily store data. Memory devices can also be used for long-term data storage. Generally, a computer also includes, or is operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, *e.g.*, magnetic, magneto-optical disks, or optical disks. A computer can also be operatively coupled to a communications network in order to receive instructions and/or data from the network and/or to transfer instructions and/or data to the network. Computer-readable storage mediums suitable for embodying computer program instructions and data include all forms of volatile and non-volatile memory, including by way of example semiconductor memory devices, *e.g.*, DRAM, SRAM, EPROM, EEPROM, and flash memory devices; magnetic disks, *e.g.*, internal hard disks or removable disks; magneto-optical disks; and optical disks, *e.g.*, CD, DVD, HD-DVD, and Blu-ray disks. The processor and the memory can be supplemented by and/or incorporated in special purpose logic circuitry.

[0087] To provide for interaction with a user, the above described techniques can be implemented on a computing device in communication with a display device, *e.g.*, a CRT

(cathode ray tube), plasma, or LCD (liquid crystal display) monitor, a mobile device display or screen, a holographic device and/or projector, for displaying information to the user and a keyboard and a pointing device, *e.g.*, a mouse, a trackball, a touchpad, or a motion sensor, by which the user can provide input to the computer (*e.g.*, interact with a user interface element). Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback, *e.g.*, visual feedback, auditory feedback, or tactile feedback; and input from the user can be received in any form, including acoustic, speech, and/or tactile input.

[0088] The above described techniques can be implemented in a distributed computing system that includes a back-end component. The back-end component can, for example, be a data server, a middleware component, and/or an application server. The above described techniques can be implemented in a distributed computing system that includes a front-end component. The front-end component can, for example, be a client computer having a graphical user interface, a Web browser through which a user can interact with an example implementation, and/or other graphical user interfaces for a transmitting device. The above described techniques can be implemented in a distributed computing system that includes any combination of such back-end, middleware, or front-end components.

[0089] The components of the computing system can be interconnected by transmission medium, which can include any form or medium of digital or analog data communication (*e.g.*, a communication network). Transmission medium can include one or more packet-based networks and/or one or more circuit-based networks in any configuration. Packet-based networks can include, for example, the Internet, a carrier internet protocol (IP) network (*e.g.*, local area network (LAN), wide area network (WAN), campus area network (CAN), metropolitan area network (MAN), home area network (HAN)), a private IP network, an IP private branch exchange (IPBX), a wireless network (*e.g.*, radio access network (RAN), Bluetooth, Wi-Fi, WiMAX, general packet radio service (GPRS) network, HiperLAN), and/or

other packet-based networks. Circuit-based networks can include, for example, the public switched telephone network (PSTN), a legacy private branch exchange (PBX), a wireless network (*e.g.*, RAN, code-division multiple access (CDMA) network, time division multiple access (TDMA) network, global system for mobile communications (GSM) network), and/or other circuit-based networks.

[0090] Information transfer over transmission medium can be based on one or more communication protocols. Communication protocols can include, for example, Ethernet protocol, Internet Protocol (IP), Voice over IP (VOIP), a Peer-to-Peer (P2P) protocol, Hypertext Transfer Protocol (HTTP), Session Initiation Protocol (SIP), H.323, Media Gateway Control Protocol (MGCP), Signaling System #7 (SS7), a Global System for Mobile Communications (GSM) protocol, a Push-to-Talk (PTT) protocol, a PTT over Cellular (POC) protocol, Universal Mobile Telecommunications System (UMTS), 3GPP Long Term Evolution (LTE) and/or other communication protocols.

[0091] Devices of the computing system can include, for example, a computer, a computer with a browser device, a telephone, an IP phone, a mobile device (*e.g.*, cellular phone, personal digital assistant (PDA) device, smart phone, tablet, laptop computer, electronic mail device), and/or other communication devices. The browser device includes, for example, a computer (*e.g.*, desktop computer and/or laptop computer) with a World Wide Web browser (*e.g.*, Chrome™ from Google, Inc., Microsoft® Internet Explorer® available from Microsoft Corporation, and/or Mozilla® Firefox available from Mozilla Corporation). Mobile computing device include, for example, a Blackberry® from Research in Motion, an iPhone® from Apple Corporation, and/or an Android™-based device. IP phones include, for example, a Cisco® Unified IP Phone 7985G and/or a Cisco® Unified Wireless Phone 7920 available from Cisco Systems, Inc.

[0092] Comprise, include, and/or plural forms of each are open ended and include the listed parts and can include additional parts that are not listed. And/or is open ended and includes one or more of the listed parts and combinations of the listed parts.

[0093] One skilled in the art will realize the invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting of the invention described herein.

CLAIMS

What is claimed is:

1. A computerized method for authenticating an end point device participating in a media-based conference call, the method comprising:

receiving, by a call processing module of a server computing device, a request to join a conference call between a plurality of end point devices, the request including a media stream associated with the first end point device;

analyzing, by the call processing module, the media stream using a matrix of authentication factors, wherein the authentication factors include at least two of: user-specific facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream;

determining, by the call processing module, an authentication score for the first end point device based upon the media stream analysis; and

determining, by the call processing module, whether to connect the first end point device to the conference call, another media resource, or another user of a network or communication system based upon the authentication score.

2. The method of claim 1, further comprising:

determining, by the call processing module, an identity of a user of the first end point device based upon the media stream analysis; and

retrieving, by the call processing module, a user profile based upon the identity of the user, the user profile including a set of permissions associated with conference call functionality.

3. The method of claim 1, wherein the acoustic environment attributes include (i) one or more pre-recorded audio files from a user and (ii) acoustic attributes of a room in which the first end point device is located.

4. The method of claim 1, wherein the visual environment attributes include (i) user-provided or system-captured still images and multiple video frames and (ii) lighting attributes of a room in which the first end point device is located.

5. The method of claim 1, wherein the technical attributes of the first end point device include an image resolution of a camera coupled to the first end point device, a device identifier, a location, and an originating address.

6. The method of claim 1, wherein the technical attributes of the media stream include a media format and a media quality.

7. The method of claim 1, further comprising:

transmitting, by the call processing module, the media stream associated with the first end point device to one or more other end point devices connected to the conference call;

receiving, by the call processing module, a validation signal from the one or more other end point devices to confirm an identity of a user of the first end point device; and

adding, by the call processing module, the received validation signal to the matrix of authentication factors.

8. The method of claim 1, wherein the media stream analysis includes comparing the matrix of authentication factors for the media stream to a matrix of authentication factors for prior media streams associated with the first end point device.

9. The method of claim 8, wherein the media stream analysis includes analyzing the matrix of authentication factors for the media stream based upon user- and/or system-specified preferences.

10. The method of claim 1, further comprising storing, by the call processing module, the media stream analysis and authentication score for the media stream in a database.

11. The method of claim 1, wherein the media stream analysis and authentication score determination occur periodically during the conference call.

12. The method of claim 1, wherein the authentication score is below a predetermined threshold, the method further comprising transmitting, by the call processing module, a request for user credentials to the first end point device.

13. The method of claim 1, further comprising transmitting, by the call processing module, the authentication score for the first end point device for display on one or more other end point devices connected to the conference call.

14. The method of claim 1, further comprising:

denying, by the call processing module, the request to join the conference call when the authentication score is below a predetermined threshold; and

disconnecting, by the call processing module, the first end point device.

15. The method of claim 1, further comprising:

allowing, by the call processing module, the request to join the conference call when the authentication score is at or above a predetermined threshold; and

connecting, by the call processing module, the first end point device to the conference call.

16. A system for authenticating an end point device participating in a media-based conference call, the system comprising a server computing device having a call processing module configured to:

receive a request to join a conference call between a plurality of end point devices, the request including a media stream associated with the first end point device;

analyze the media stream using a matrix of authentication factors, wherein the authentication factors include at least two of: user-specific facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream;

determine an authentication score for the first end point device based upon the media stream analysis; and

determine whether to connect the first end point device to the conference call, another media resource, or another user of a network or communication system based upon the authentication score.

17. The system of claim 16, wherein the call processing module is further configured to:

determine an identity of a user of the first end point device based upon the media stream analysis; and

retrieve a user profile based upon the identity of the user, the user profile including a set of permissions associated with conference call functionality.

18. The system of claim 16, wherein the acoustic environment attributes include (i) one or more pre-recorded audio files from a user and (ii) acoustic attributes of a room in which the first end point device is located.

19. The system of claim 16, wherein the visual environment attributes include (i) user-provided or system-captured still images and multiple video frames and (ii) lighting attributes of a room in which the first end point device is located.

20. The system of claim 16, wherein the technical attributes of the first end point device include an image resolution of a camera coupled to the first end point device, a device identifier, a location, and an originating address.

21. The system of claim 16, wherein the technical attributes of the media stream include a media format and a media quality.

22. The system of claim 16, wherein the call processing module is further configured to:

transmit the media stream associated with the first end point device to one or more other end point devices connected to the conference call;

receive a validation signal from the one or more other end point devices to confirm an identity of a user of the first end point device; and

add the received validation signal to the matrix of authentication factors.

23. The system of claim 16, wherein the media stream analysis includes comparing the matrix of authentication factors for the media stream to a matrix of authentication factors for prior media streams associated with the first end point device.

24. The method of claim 23, wherein the media stream analysis includes analyzing the matrix of authentication factors for the media stream based upon user- and/or system-specified preferences.

25. The system of claim 16, wherein the call processing module is further configured to store the media stream analysis and authentication score for the media stream in a database.

26. The system of claim 16, wherein the media stream analysis and authentication score determination occur periodically during the conference call.

27. The system of claim 16, wherein the authentication score is below a predetermined threshold, the call processing module further configured to transmit a request for user credentials to the first end point device.

28. The system of claim 16, wherein the call processing module is further configured to:

deny the request to join the conference call when the authentication score is below a predetermined threshold; and

disconnect the first end point device.

29. The system of claim 16, wherein the call processing module is further configured to:

allow the request to join the conference call when the authentication score is at or above a predetermined threshold; and

connect the first end point device to the conference call.

30. A computer program product, tangibly embodied in a non-transitory computer readable storage device, for authenticating an end point device participating in a media-based conference call, the computer program product including instructions operable to cause a call processing module of a server computing device to:

receive a request to join a conference call between a plurality of end point devices, the request including a media stream associated with the first end point device;

analyze the media stream using a matrix of authentication factors, wherein the authentication factors include at least two of: user-specific facial recognition attributes, user-specific audio recognition attributes, acoustic environment attributes, visual environment attributes, user gesture attributes, technical attributes of the end point device, and technical attributes of the media stream;

determine an authentication score for the first end point device based upon the media stream analysis; and

determine whether to connect the first end point device to the conference call, another media resource, or another user of a network or communication system based upon the authentication score.

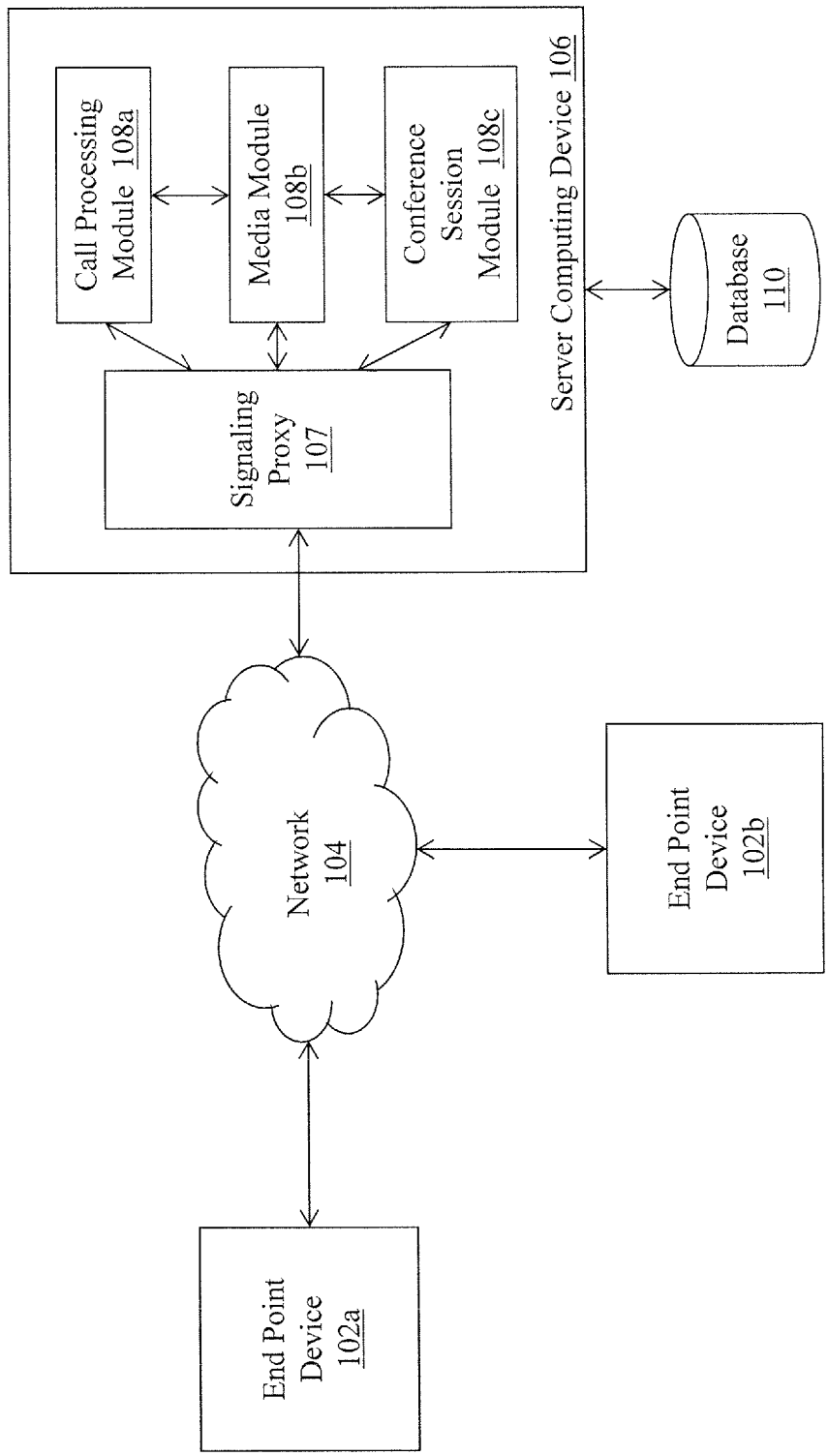
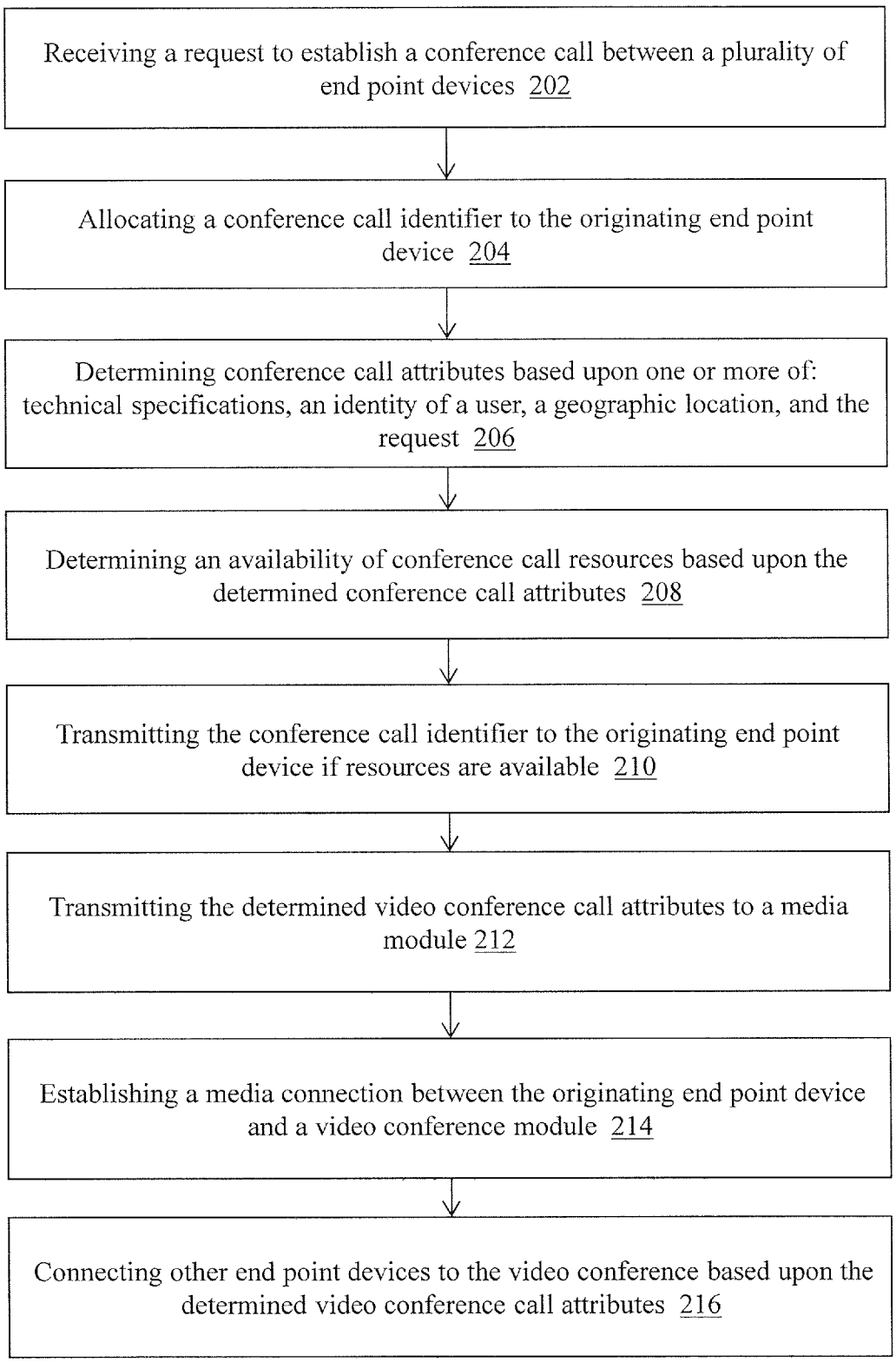

Server Computing Device 106
Call Processing Module 108a
Media Module 108b
Conference Session Module 108c
Signaling Proxy 107
Database 110
Network 104
End Point Device 102a
End Point Device 102b
100

FIG. 1


Receiving a request to establish a conference call between a plurality of end point devices 202
Allocating a conference call identifier to the originating end point device 204
Determining conference call attributes based upon one or more of: technical specifications, an identity of a user, a geographic location, and the request 206
Determining an availability of conference call resources based upon the determined conference call attributes 208
Transmitting the conference call identifier to the originating end point device if resources are available 210
Transmitting the determined video conference call attributes to a media module 212
Establishing a media connection between the originating end point device and a video conference module 214
Connecting other end point devices to the video conference based upon the determined video conference call attributes 216

200

FIG. 2

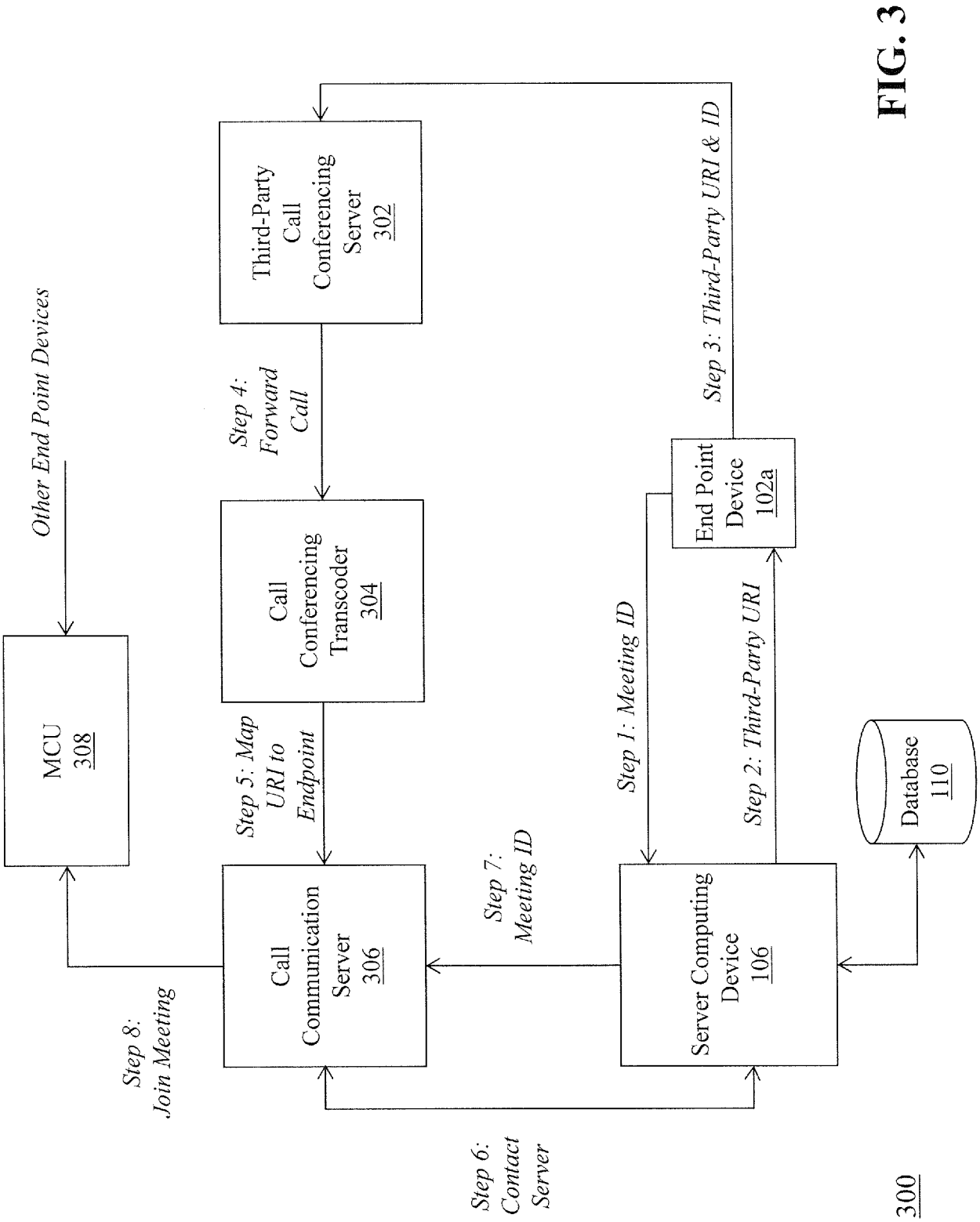

Other End Point Devices
MCU
308
Third-Party
Call
Conferencing
Server
302
Step 4:
Forward
Call
Call
Conferencing
Transcoder
304
Step 5: Map
URI to
Endpoint
Step 3: Third-Party URI & ID
End Point
Device
102a
Step 1: Meeting ID
Step 2: Third-Party URI
Database
110
Call
Communication
Server
306
Step 7:
Meeting ID
Server Computing
Device
106
Step 8:
Join Meeting
Step 6:
Contact
Server
300

FIG. 3

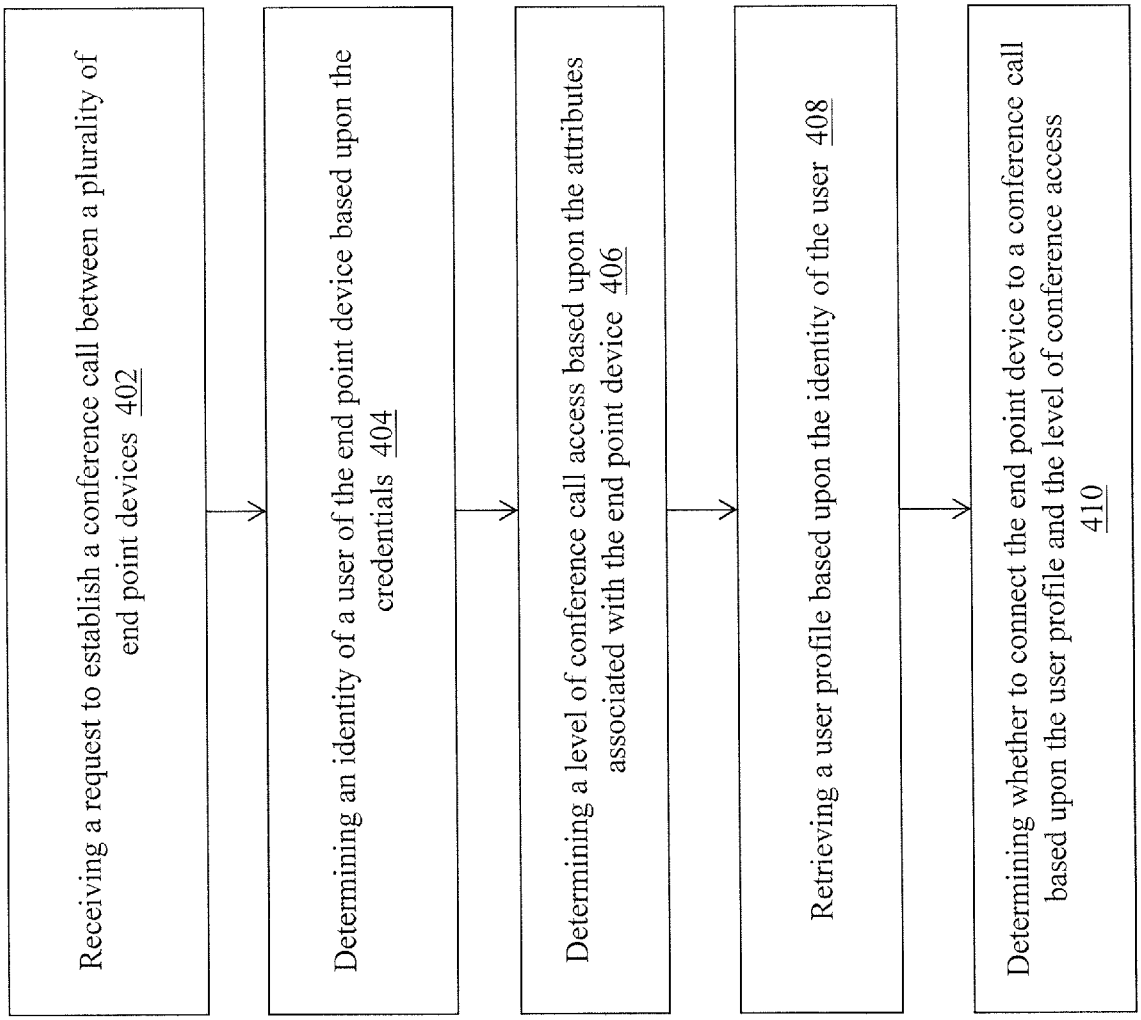

Receiving a request to establish a conference call between a plurality of end point devices 402
Determining an identity of a user of the end point device based upon the credentials 404
Determining a level of conference call access based upon the attributes associated with the end point device 406
Retrieving a user profile based upon the identity of the user 408
Determining whether to connect the end point device to a conference call based upon the user profile and the level of conference access 410
400

FIG. 4

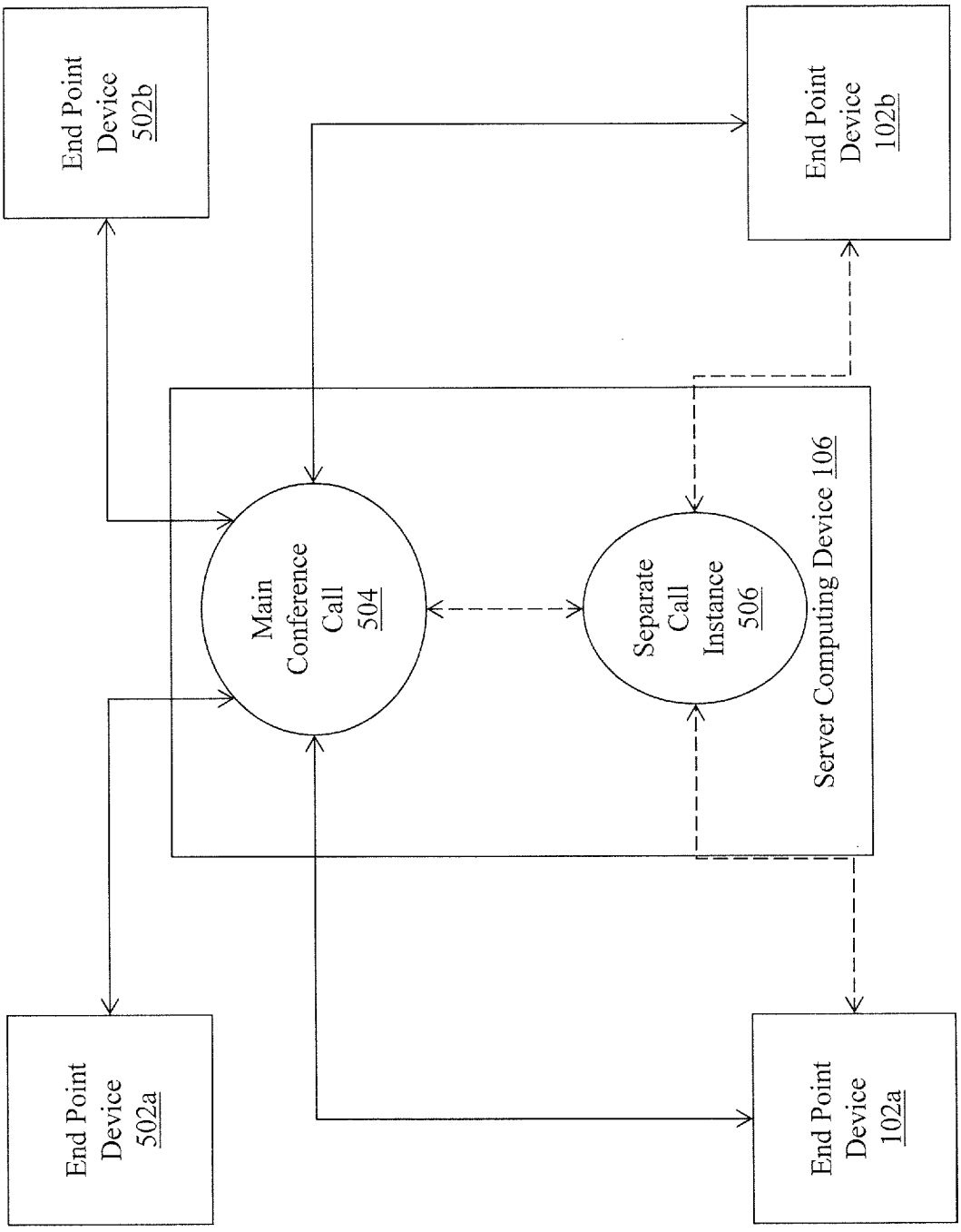

End Point Device 502b
End Point Device 102b
Main Conference Call 504
Separate Call Instance 506
Server Computing Device 106
End Point Device 502a
End Point Device 102a
500

FIG. 5

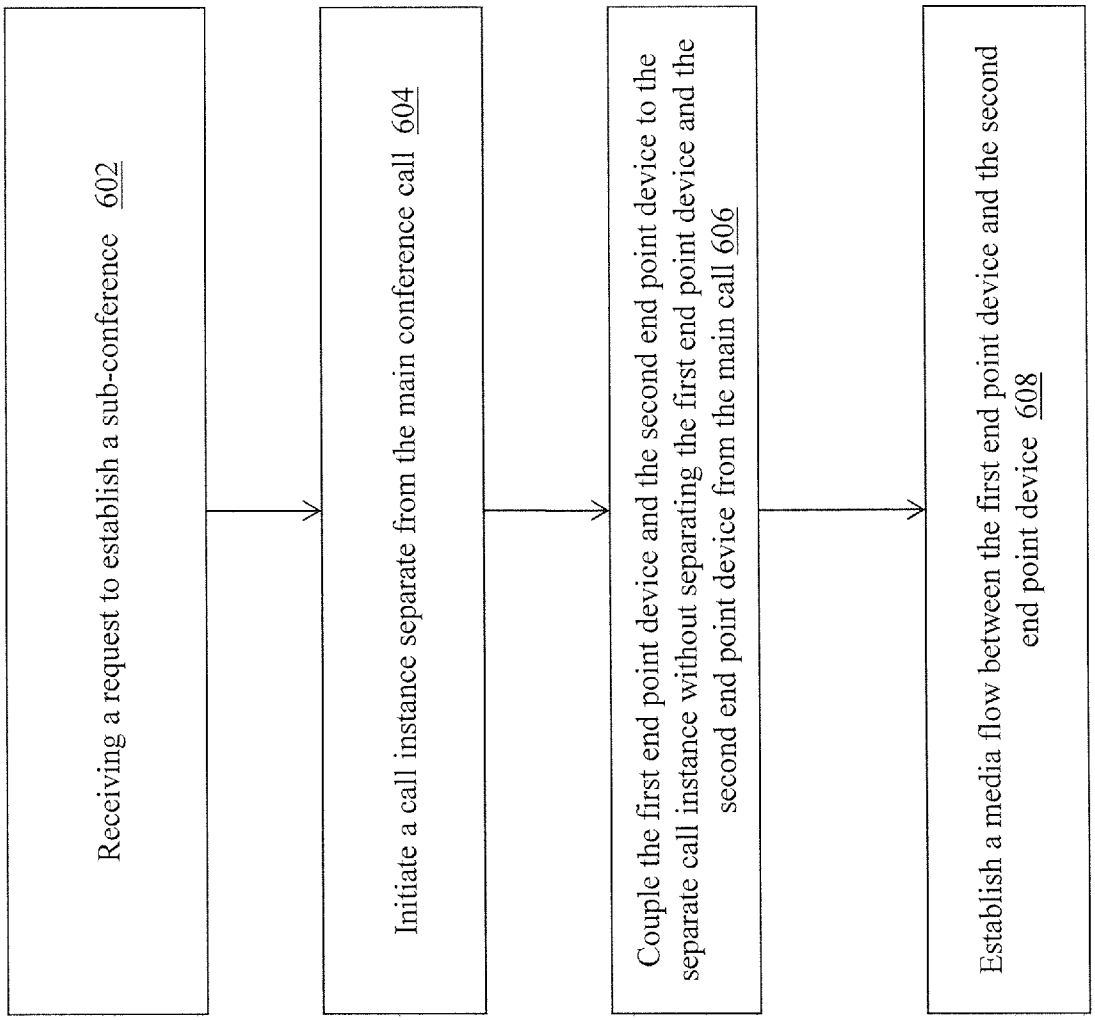

Receiving a request to establish a sub-conference 602
Initiate a call instance separate from the main conference call 604
Couple the first end point device and the second end point device to the separate call instance without separating the first end point device and the second end point device from the main call 606
Establish a media flow between the first end point device and the second end point device 608
600

FIG. 6

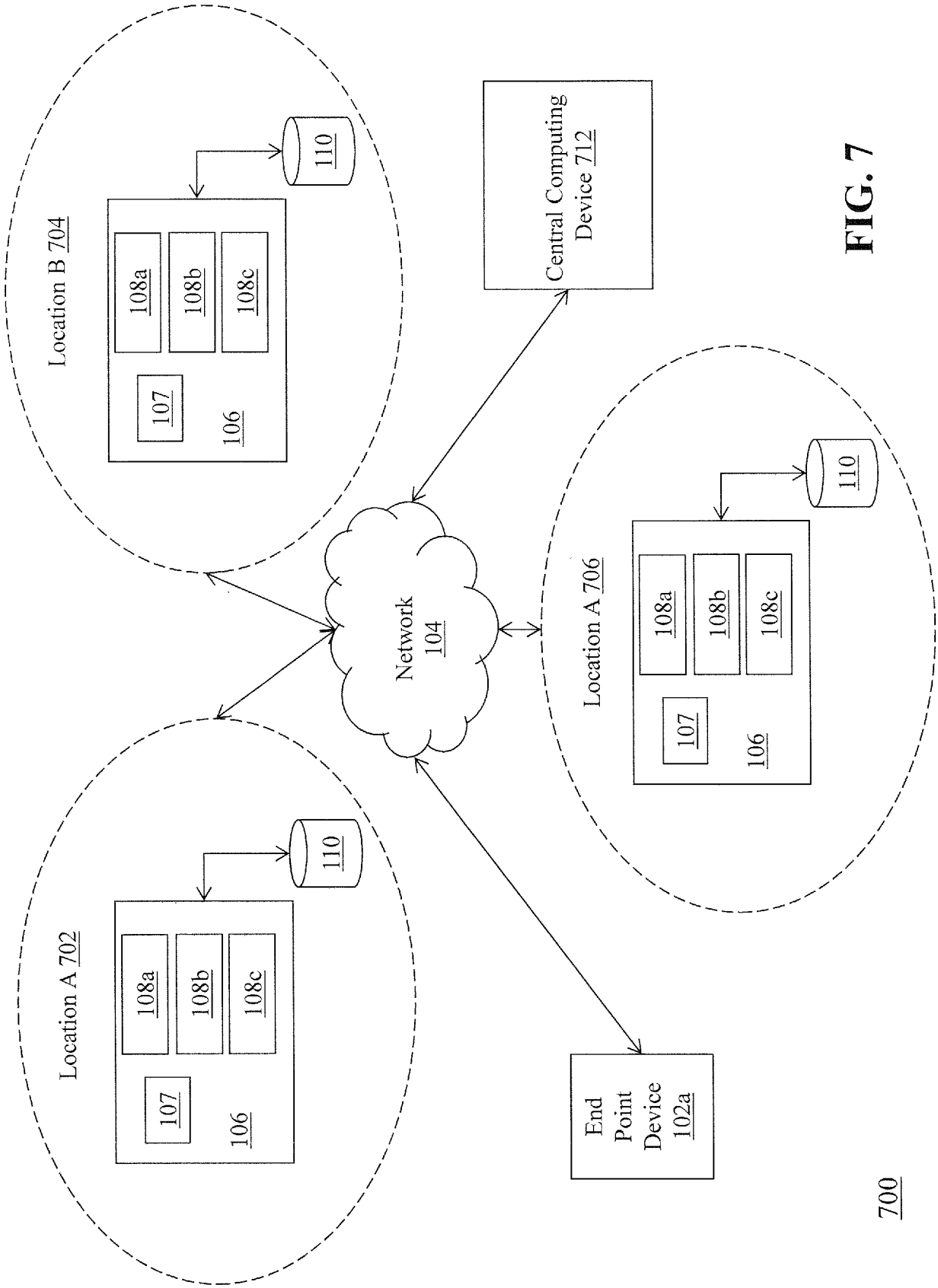

Location B 704
110
108a
108b
108c
107
106
Central Computing Device 712
Network 104
Location A 706
110
108a
108b
108c
107
106
Location A 702
110
108a
108b
108c
107
106
End Point Device 102a
700

FIG. 7

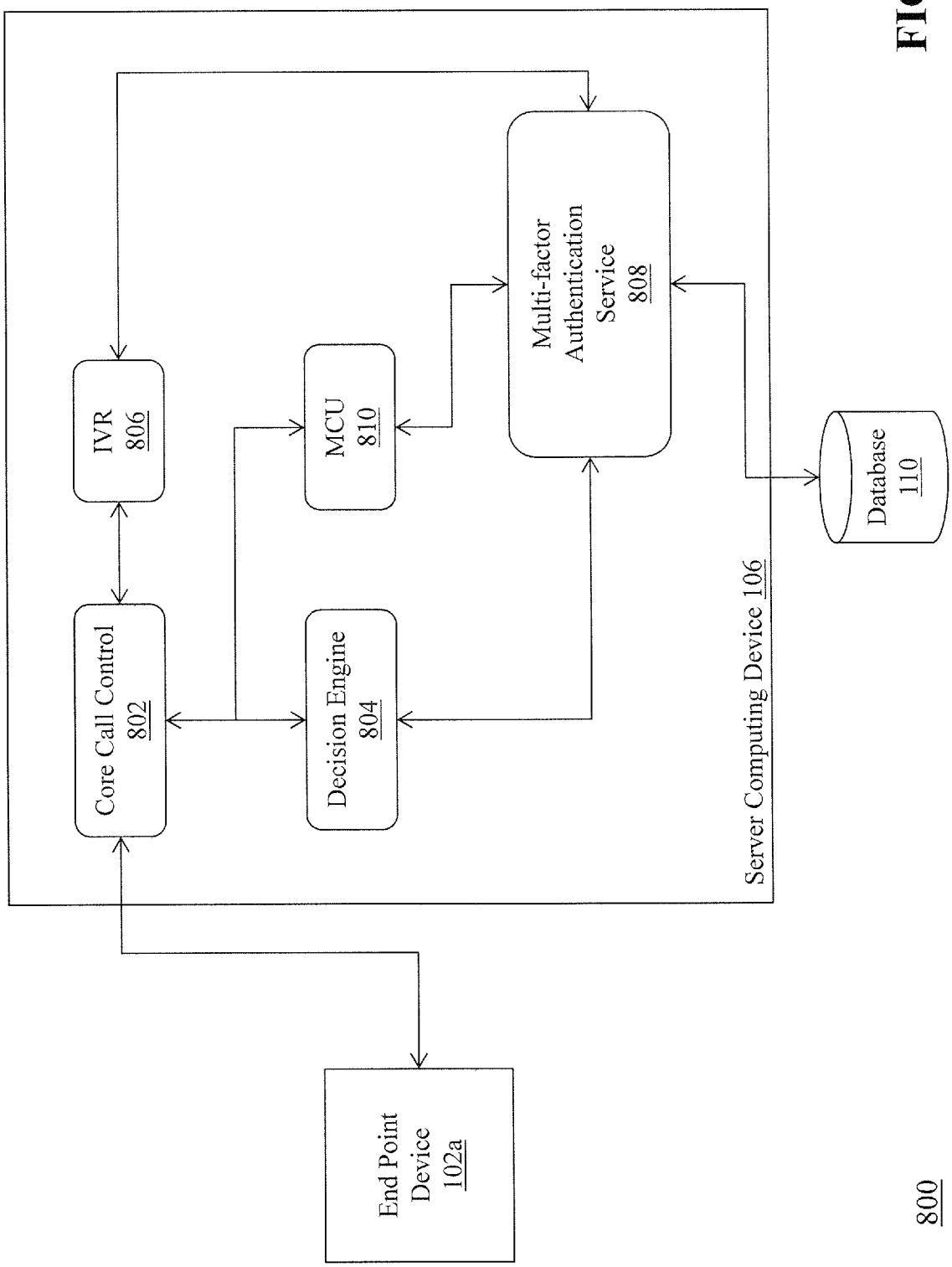

Server Computing Device 106
IVR 806
Core Call Control 802
MCU 810
Decision Engine 804
Multi-factor Authentication Service 808
Database 110
End Point Device 102a
800

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US15/49011

A. CLASSIFICATION OF SUBJECT MATTER
IPC: **H04N 7/14**(2006.01)

USPC: 348/14.01-14.16
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 348/14.01-14.16

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
US-PGPUB, USPAT, FPRS, EPO, JPO: screened, face, display, authentication, score, displaying, conference, displayed

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2006/0259755 A1 (KENOYER) 16 November 2006 (16.11.2006), see entire documents.	1-7, 8-30
Y	US 2014/0289834 A1 (LINDERMANN) 25 September 2014 (25.09.2014), see entire documents.	1-7, 8-30
Y	US 2014/0111596 A1 (GREVERS, JR) 24 April 2014 (24.04.2014), see entire documents.	7, 22
Y	US 2013/0329970 A1 (IRIE et al) 12 December 2013 (12.12.2013), see entire documents.	13
A	US 2013/0133049 A1 (PEIRCE) 23 May 2013 (23.05.2013), see entire documents.	1-30

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

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
09 October 2015 (09.10.2015)	15 OCT 2015
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer William Krynski Telephone No. 571-272-1700