



(51) International Patent Classification:

H04N 7/18 (2006.01) *H04N 21/647* (2011.01)
H04N 21/2343 (2011.01)

(21) International Application Number:

PCT/IB2023/055782

(22) International Filing Date:

05 June 2023 (05.06.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant: **TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)** [SE/SE]; SE-164 83, Stockholm (SE).

(72) Inventors: **SZABÓ, Géza**; Fenyves u. 5, 6000 Kecskemet (HU). **LAKI, Sándor**; Matróz utca 6. 9/24, 1033 Budapest

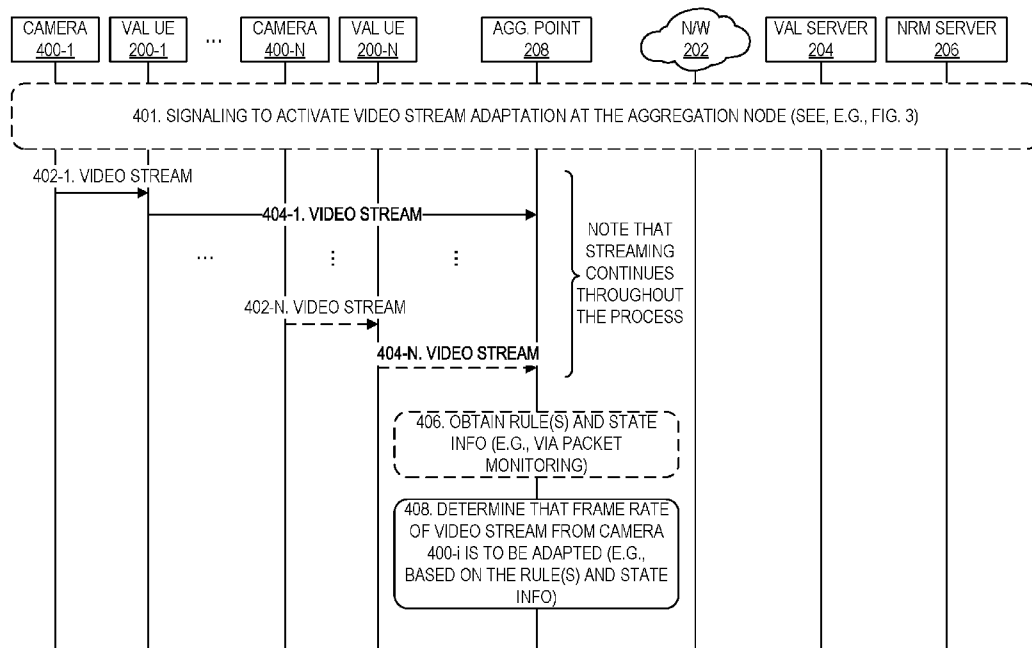
(HU). **GYÖRGYI, Csaba**; Puttony utca 38., H-4405 Nyíregyháza (HU). **VÖRÖS, Péter**; Andor utca 6.8 em 36, H-1119 Budapest (HU). **KECSKEMÉTI, Károly**; Budai nagy Antal u.9, HU-6200 Kiskörös (HU).

(74) Agent: **BEVINS, R. Chad**; 106 Pinedale Springs Way, Cary, North Carolina 27511 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH,

(54) Title: IN-NETWORK ADAPTIVE QUALITY CONTROL OF IP CAMERAS

FIG. 4A



(57) Abstract: Systems and methods are disclosed for in-network adaptive quality control of video streams from cameras. In one embodiment, a method of operation of radio node for in-network adaptive quality control comprises, for a streaming video camera from among one or more streaming video cameras associated to the radio node, deciding to adapt a frame rate of a video stream received from the streaming video camera. The method further comprises, in response to deciding to adapt the frame rate of the video stream received from the streaming video camera, receiving the video stream from the stream video camera wherein the video stream has a first frame rate, adapting a frame rate of the video stream to provide an adapted video stream having a second frame rate that is lower than the first frame rate, and transmitting the adapted video stream at the second frame rate via a wireless network.

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

IN-NETWORK ADAPTIVE QUALITY CONTROL OF IP CAMERASTechnical Field

[0001] The present disclosure relates to in-network adaptive quality control for a
5 video stream(s) from a camera(s).

Background

[0002] Third Generation Partnership Project (3GPP) Technical Specification (TS)
23.434 (see, e.g., V18.4.1) defines a functional architecture for Service Enabler
10 Architecture Layer (SEAL) over 3GPP networks to support vertical applications such as,
e.g., Vehicle to Anything (V2X) applications. As part of this functional architecture,
3GPP 23.434 defines Network Resource Management (NRM) functionality which is
implemented on the radio channel. While the NRM affects the radio channel, the Vertical
Application Layer (VAL) application data transferred over the radio channel is not altered
15 by the 3GPP network.

[0003] One example of a vertical application that can transport data via a 3GPP
network using the functional architecture for SEAL over 3GPP networks defined in 3GPP
TS 23.434 is a vertical application for an environment (e.g., an industrial environment
that employs robotics) that uses video streams from multiple (e.g., many) cameras
20 located in the environment. For example, in industrial environments, video streams from
cameras located in the industrial environment are needed for certain tasks (e.g., pick
and place, etc.) to control and observe the setup.

[0004] There is a need for systems and methods for providing NRM for vertical
applications that transport video streams from multiple cameras via the 3GPP network.
25

Summary

[0005] Systems and methods are disclosed for in-network adaptive quality control of
video streams from cameras. In one embodiment, a method of operation of radio node
for in-network adaptive quality control of video streams from streaming video cameras
30 comprises, for a streaming video camera from among one or more streaming video
cameras associated to the radio node, deciding to adapt a frame rate of a video stream
received from the streaming video camera. The method further comprises, in response
to deciding to adapt the frame rate of the video stream received from the streaming

video camera, receiving the video stream from the stream video camera wherein the video stream has a first frame rate, adapting a frame rate of the video stream to provide an adapted video stream having a second frame rate that is lower than the first frame rate, and transmitting the adapted video stream at the second frame rate via a wireless network. In this manner, outgoing video traffic transported through the network may be reduced.

[0006] In one embodiment, the one or more streaming video cameras associated to the device comprise a plurality of streaming video cameras, the radio node is a network node of the wireless network, and the radio node operates as an aggregation point for video streams from the plurality of streaming video cameras. In one embodiment, the network node is a network switch.

[0007] In one embodiment, the radio node is a wireless communication device enabled to transmit and receive wireless signals to and from the wireless network.

[0008] In one embodiment, the wireless network is a Radio Access Network (RAN) of a cellular communications system, the radio node is a radio node of the cellular communications system, and the method further comprises receiving a request to perform video stream adaptation from a Network Resource Management (NRM) server or an NRM client associated to the cellular communications system. Deciding to adapt the frame rate of the video stream received from the streaming video camera is

responsive to receiving the request. In one embodiment, the radio node is a network node associated to a RAN of the cellular communications system. In one embodiment, the one or more streaming video cameras associated to the device comprise a plurality of streaming video cameras, and the radio node operates as a network-controlled aggregation point for video streams from the plurality of streaming video cameras. In one embodiment, the network node is a network switch.

[0009] In one embodiment, the radio node is a User Equipment (UE) enabled to transmit and receive wireless signals to and from a RAN of the cellular communications system.

[0010] In one embodiment, the video stream comprises both frames of a first frame type and frames of a second frame type that are dependent, directly or indirectly, on the frames of the first frame type, wherein adapting the frame rate of the video stream comprises dropping at least some of the frames of the second frame type. In one embodiment, the frames of the first frame type are key frames, and the frames of the

second frame type are predicted frames. In one embodiment, the video stream is an Internet Protocol (IP) stream comprising a plurality of IP packets, and dropping at least some of the frames of the second frame type comprises dropping a subset of the plurality of packets that carry the at least some of the frames of the second frame type.

- 5 In one embodiment, dropping at least some of the frames of the second frame type comprises dropping all frames of the second frame type. In another embodiment, dropping at least some of the frames of the second frame type comprises dropping only some frames of the second frame type. In another embodiment, the method further comprises determining an amount of the frames of the second frame type to be
- 10 dropped, wherein dropping at least some of the frames of the second frame type comprises dropping the determined amount of the frames of the second frame type.

- [0011]** In one embodiment, the video stream comprises both frames of a first frame type and frames of a second frame type that are dependent, directly or indirectly, on the frames of the first frame type, wherein adapting the frame rate of the video stream
- 15 comprises dropping both all frames of the first frame type and all frames of the second frame type. In one embodiment, the frames of the first frame type are key frames, and the frames of the second frame type are predicted frames.

- [0012]** In one embodiment, deciding to adapt the frame rate of the video stream received from the streaming video camera comprises deciding to adapt the frame rate
- 20 of the video stream received from the streaming video camera based on one or more predefined rules and state information of one or more secondary devices located with an environment captured by the one or more streaming video cameras. In one embodiment, the one or more secondary devices move within the environment captured by the one or more streaming video cameras. In one embodiment, the one or more
- 25 secondary devices are robotic devices that move within the environment captured by the one or more streaming video cameras. In another embodiment, the state information of the one or more secondary devices comprises information that indicates locations of the one or more secondary devices within the environment. In another embodiment, the one or more secondary devices comprise one or more sensors within
- 30 the environment captured by the one or more stream video cameras. In one embodiment, the method further comprises obtaining the state information of the one or more secondary devices by monitoring packets transmitted from the one or more secondary devices to an application server via the radio node.

[0013] In one embodiment, the method further comprises repeating the method for one or more additional stream video cameras from among the one or more streaming video cameras associated to the radio node.

[0014] In one embodiment, the method further comprises deciding to deactivate
5 adaptation of the frame rate of the video stream received from the streaming video camera and, in response to deciding to deactivate adaptation of the frame rate of the video stream received from the streaming video camera, stopping the adapting of the frame rate of the video stream starting at a certain frame in the video stream. In one
10 embodiment, the certain frame is a first key frame of the video stream that occurs after deciding to deactivate adaptation of the frame rate of the video stream.

[0015] Corresponding embodiments of a radio node are also disclosed. In one embodiment, a radio node for in-network adaptive quality control of video streams from streaming video cameras comprises processing circuitry and memory comprising
15 instructions executable by the processing circuitry whereby the radio node is caused to, for a streaming video camera from among one or more streaming video cameras associated to the radio node, decide to adapt a frame rate of a video stream received from the streaming video camera. The radio node is further caused to, in response to
20 deciding to adapt the frame rate of the video stream received from the streaming video camera, receive the video stream from the stream video camera, the video stream having a first frame rate, adapt a frame rate of the video stream to provide an adapted video stream having a second frame rate that is lower than the first frame rate, and
transmit the adapted video stream at the second frame rate via a wireless network.

Brief Description of the Drawings

25 **[0016]** The accompanying drawing figures incorporated in and forming a part of this specification illustrate several aspects of the disclosure, and together with the description serve to explain the principles of the disclosure.

[0017] Figure 1 illustrates one example system in which embodiments of the present disclosure may be implemented;

30 **[0018]** Figure 2 illustrates modified version of the architecture defined by 3rd Generation Partnership Project (3GPP) Technical Specification (TS) 23.434 in accordance with an embodiment of the present disclosure;

[0019] Figure 3 illustrates the operation of the system of Figure 2 in accordance with one embodiment of the present disclosure;

[0020] Figures 4A through 4C illustrate the operation of the system of Figure 2 in accordance with some embodiments of the present disclosure;

5 **[0021]** Figure 5 is a flow chart that illustrates the operation of the aggregation point to adapt video streams in accordance with one example embodiment of the present disclosure;

[0022] Figure 6 is a flow chart that illustrates the operation of the aggregation point to adapt video streams in accordance with another example embodiment of the present
10 disclosure;

[0023] Figure 7 illustrates the operation of the system of Figure 2 in accordance with another embodiment of the present disclosure;

[0024] Figures 8A and 8B illustrate the operation of the system of Figure 2 in accordance with some embodiments of the present disclosure;

15 **[0025]** Figures 9 and 10 are schematic block diagrams of example embodiments of a network node in which aspects of the present disclosure may be implemented; and

[0026] Figures 11 and 12 are schematic block diagrams of example embodiments of a User Equipment (UE) in which aspects of the present disclosure may be implemented.

20 Detailed Description

[0027] The embodiments set forth below represent information to enable those skilled in the art to practice the embodiments and illustrate the best mode of practicing the embodiments. Upon reading the following description in light of the accompanying drawing figures, those skilled in the art will understand the concepts of the disclosure
25 and will recognize applications of these concepts not particularly addressed herein. It should be understood that these concepts and applications fall within the scope of the disclosure.

[0028] Note that the description given herein focuses on a 3rd Generation Partnership Project (3GPP) cellular communications system and, as such, 3GPP terminology or
30 terminology similar to 3GPP terminology is oftentimes used. However, the concepts disclosed herein are not limited to a 3GPP system.

[0029] As discussed above, 3GPP Technical Specification (TS) 23.434 (see, e.g., V18.4.1) defines a functional architecture for Service Enabler Architecture Layer (SEAL)

over 3GPP networks to support vertical applications such as, e.g., Vehicle to Anything (V2X) applications. As part of this functional architecture, 3GPP 23.434 defines Network Resource Management (NRM) functionality which is implemented on the radio channel. While the NRM affects the radio channel, the Vertical Application Layer (VAL) application data transferred over the radio channel is not altered by the 3GPP network.

[0030] One example of a vertical application that can transport data via a 3GPP network using the functional architecture for SEAL over 3GPP networks defined in 3GPP TS 23.434 is a vertical application for an environment (e.g., an industrial environment that employs robotics) that uses video streams from multiple (e.g., many) cameras located in the environment. For example, in industrial environments, video streams from cameras located in the industrial environment are needed for certain tasks (e.g., pick and place, etc.) to control and observe the setup. However, not all the video streams are needed at high Frames-Per-Second (FPS) all the time. A robot may be at different locations, in different states, etc., determining which streams are needed.

[0031] Video streams usually contain key frames (oftentimes referred to as "Intra-coded frames" or "I-frames") describing a full image at a given moment in time. Intermediate moments can be assembled using differences compared to these key frames. For example, so called "Predicted frames" or "P-frames" reuse information from the past frames (e.g., past I-frame(s) and/or past P-frame(s)). In other words, P-frames can be anchored to either I-frames or other P-frames. When transferring a video stream via Internet Protocol (IP), one frame can occupy multiple IP packets.

[0032] If the 3GPP NRM functionally cannot request that the VAL layer adapt the VAL layer traffic and the VAL traffic is not self-adapting, then the VAL traffic can suffer significant performance degradation. If the VAL traffic includes a video stream, the video stream may be degraded to the point that it cannot be decoded by the receiver.

[0033] When VAL traffic includes video streams from multiple cameras, maintaining high-resolution and high Frames-Per-Second (FPS) video streams for all of the cameras would put a huge load on the 3GPP network and on the receiving device(s) (e.g., server(s)). Moreover, multiple high-resolution and high-FPS video streams can be costly for certain cloud services where the network traffic is also billed along with other services, or in cases where the aggregated traffic needs to be transported across a link with limited capacity.

[0034] Existing adaptive solutions (e.g., the Open Network Video Interface Forum (ONVIF) solution) require special capabilities and cooperation both from the sender and the receiver of the video streams.

[0035] Further, modifying the video stream settings at the cameras in real-time often results in a large reconfiguration delay that is unacceptable for many VAL applications.

[0036] Systems and methods are disclosed herein that address the aforementioned and/or other challenges. Embodiments of systems and methods are disclosed that enable network-controlled adaptive quality control of video streams. In other words, the NRM functionality of the network enabled to adaptively control the quality of the video streams, e.g., such that the video streams (i.e., the VAL traffic) is fit into the radio channel. In one embodiment, an aggregation point of the video streams is network-controlled and operates to aggregate the video streams into a combined VAL traffic flow that is then transported by the radio channel provided by the 3GPP network. This aggregation point may, for example, be implemented in a UE of the 3GPP network or a transport layer of the 3GPP network. The aggregation point operates to provide network-controlled adaptive quality control of the video streams by, e.g., dropping packets that carry at least some frames (e.g., some or all P-frames) of the video streams subject to certain rules or conditions.

[0037] In one embodiment, the aggregation point is a programmable router. A programmable router can provide much more functionality than pure packet forwarding. A programmable router can read fields included in the packets, e.g., for high-throughput calculations on the application level during communication.

[0038] In one embodiment, the aggregation point includes programmable data planes that observe both a state of an environment (e.g., an industrial environment) in which the cameras are located and the video streams. The aggregation point then adapts one or more of the video streams based on the state of the environment. In one embodiment, the adaptation is performed by filtering or dropping packets carrying information for a selected set of frames (e.g., P-frames in H264 and H265) from the video stream(s) without compromising the connection. In one embodiment, this filtering is dynamically turned on and off based on the observed state of the environment and one or more associated rules (e.g., Feed A is always filtered except when robot X is within a given area). For example, in an industrial environment in which robotic devices move around the environment, a rule may be defined that a video stream from any

camera that captures video for an area in which no robotic device is located is to be adapted to reduce quality by, e.g., filtering or dropping all packets that carry frames of that video stream(s), filtering or dropping all packets that carry a certain type of frame (e.g., P-frames, e.g., in H264 and H265) of that video stream(s), or filtering or dropping
5 some of the frames of that video stream(s) (e.g., drop a subset of, or certain amount of, P-frames of the video stream(s)).

[0039] The observation and filtering happen inside the programmable data plane thus providing ultra-low latency and high throughput. Moreover, in one embodiment, the filtering is performed on a local network (i.e., a local network of the cameras), and
10 thus outgoing traffic (i.e., outgoing traffic transported via the 3GPP network to the edge or remote cloud for further processing) is reduced.

[0040] The proposed video stream quality control approach is independent of used cameras and their capabilities and is provided as a service by the network.

[0041] In one embodiment, new elements and interfaces are added to the functional
15 model of the 3GPP TS 23.434. In one embodiment, the existing NRM request Application Programming Interface (API) is extended with new information elements.

[0042] In one embodiment, the quality of the video stream(s) is modified in the network without requiring any support and modification by the end-nodes (i.e., both camera and the receiver side). In one embodiment, the stream quality is set according
20 to the state of the environment (e.g., robot states, robot locations, etc.).

[0043] In one embodiment, a digital packet processing circuit (e.g., in an aggregation point or in a UE) observes the state of the environment and adapts (e.g., reduces or increases) the quality of video streams for selected set of cameras. In one embodiment, the digital packet processing circuit filters all or selected frames (e.g., all
25 or selected P-frames) from the video stream(s) after quality reduction is activated, with immediate effect on the traffic stream. In one embodiment, the digital packet processing circuit deactivates filtering (e.g., deactivates P-frame filtering) when a high-quality stream is needed or desired in the environment (e.g., as determined by one or more rules). The high-quality stream is provided either immediately after the
30 deactivation of filtering or after the next I-frame sent in the video stream, depending on the P-frame anchors. In one embodiment, the digital packet processing circuit maintains a single register cell (e.g., bit, byte, etc.) for each camera to maintain if a P-frame starts or ends. The register cell is then used to identify if the packet under processing belongs

to a P-frame or not. Note a frame (e.g., both P-frame and I-frame) is generally fragmented to multiple packets. The packets are preferably Internet Protocol (IP) packets.

[0044] While not being limited to or by any particular advantage, embodiments of the present disclosure may provide a number of advantages over existing solutions. For example, embodiments of the present disclosure may provide any one or more of the following advantages:

- reduced outgoing video traffic,
- lower ingress cost for certain cloud services to which the video streams are sent,
- 10 • reduced load on receiving servers,
- does not require any changes in the existing protocol used by the devices,
- the in-network computing component can be implemented in P4,
- using the system (e.g., defining rule(s) for activating or deactivating filtering) does not require deep networking expertise, and
- 15 • parameters of the filtering rules can be changed during runtime.

Note that "P4" is a programming language for controlling packet forwarding planes in networking devices, such as routers and switches. P4 is to be distinguished from a general purpose language such as C or Python. Rather, P4 is a domain-specific language with a number of constructs optimized for network data forwarding.

- 20 **[0045]** Figure 1 illustrates one example system 100 in which embodiments of the present disclosure may be implemented. As illustrated, the system 100 includes multiple cameras 102-1 through 102-6 (e.g., IP cameras) that output continuous video streams for certain areas 104-1 to 104-4 within an environment 106 (e.g., an industrial environment such as, e.g., a manufacturing facility, a warehouse, or the like). In this
- 25 example, the camera 102-1 captures video for the area 104-3 and outputs a respective high quality (i.e., high FPS rate in the illustrated example) video stream. Likewise, the camera 102-2 captures video for the combined area of areas 104-1 and 104-3 and outputs a respective high quality (i.e., high FPS rate in the illustrated example) video stream. The camera 102-3 captures video for the area 104-1 and outputs a respective
- 30 high quality (i.e., high FPS rate in the illustrated example) video stream. The camera 102-4 captures video for the area 104-2 and outputs a respective high quality (i.e., high FPS rate in the illustrated example) video stream. The camera 102-5 captures video for the area 104-2 and outputs a respective high quality (i.e., high FPS rate in the

illustrated example) video stream. The camera 102-6 captures video for the combined area of areas 104-2 and 104-4 and outputs a respective high quality (i.e., high FPS rate in the illustrated example) video stream. In this example, a robot 108 moves within the environment 106 and corresponding state information (e.g., location of the robot within the environment 106) is transmitted by the robot 108, e.g., via status messages transmitted by the robot 108 to, e.g., cloud server(s) 114.

[0046] The system 100 also includes a network-controlled device 110 that provides connectivity to a 3GPP network 112 (e.g., a 5G or 6G 3GPP network). The network-controlled device 110 is, in one embodiment, a P4 switch, but is not limited thereto. For example, in another embodiment, the network-controlled device 110 is a User Equipment (UE) connected to the 3GPP network 112. The network-controlled device 110 provides programmable data planes for the cameras 102-1 through 102-6. In this embodiment, the video streams from the cameras 102-1 through 102-6 as well as the state information (e.g., status messages including the state of the robot 108) from the robot 108 are provided to the network-controlled device 110 for transport to, in this example, a cloud server(s) 114 (i.e., a server(s) that provides a cloud service such as, e.g., a service that monitors or controls one or more aspects of the environment 106). The cloud server(s) 114 may be a public server(s), a private server(s), or an edge server of the 3GPP network 112 that receives the video streams from the cameras 102-1 through 102-6 and the status messages from the robot 108.

[0047] In one example embodiment, the cameras 102-1 through 102-6 are IP cameras, and the video streams are transported via IP packets sent by the cameras 102-1 through 102-6 to an intended recipient, which in the illustrated example is the cloud server(s) 114. More specifically, each video streams includes a series of frames (e.g., I-frames and P-frames) where each frame is carried by one or more IP packets transmitted from the respective camera 102 to the cloud server(s) 114. Likewise, in one embodiment, the state information of the robot 108 is included in one or more IP packets that are, in this example, transmitted from the robot 108 to the cloud server(s) 114 via the network-controlled device 110 (and thus the 3GPP network 112).

[0048] The network-controlled device 110 includes a camera stream shaping function 116 that adaptively controls a quality of the video streams from the cameras 102-1 through 102-6 based on a state of the environment 106 (e.g., the state of the robot 108 in the illustrated example) and one or more user-defined filtering rules. In this example,

the filtering rules are defined or received from a user(s) (e.g., a person associated to the environment 106 such as, e.g., a representative of a company or entity that owns the environment) via a camera stream shaping Application Programming Interface (API) 118. The filtering rules are, in one embodiments, rules that check one or more certain conditions related to the state of the environment 106 and define one or more filtering related parameters to manage the quality of one or more of the video streams if the one or more certain conditions are satisfied (e.g., always filter the video stream from camera 102-2 except when the robot 108 is within area 104-4). In one example embodiment, a rule is defined for a certain camera 102-i (where "i" is, in this example, from the set {1, 2, ..., 6}) that defines the quality of the video stream (e.g., low-FPS or high-FPS) to be provided as a function of the state (e.g., location) of the robot 108. In one embodiment, the filtering rules are stateful, where robots (e.g., in the case of multiple robots 108 in the environment 106) can reach either other's states.

[0049] In operation, for each camera 102-i, the network-controlled device 110 receives the respective video stream (e.g., carried by IP packets) from the camera 102-i, dynamically adjusts (e.g., decreases or increases) the quality of the video stream based on the user-defined rule(s) and the state of the environment 106 (e.g., as observed from status messages transmitted by the robot 108), and transmits the resulting (adjusted) video stream to, in this example, the cloud server 114 via the 3GPP network 112. Note that the state of the environment 106 (e.g., the location of the robot 108) and the user-defined rule(s) can change over time and, in response, the network-controlled device 110 will change (dynamically) the adjusted quality of the video streams accordingly.

[0050] As discussed below, in one embodiment, the camera stream shaping function 116 of the network-controlled device 110 adapts the quality of a video stream by filtering (i.e., dropping) at least some of the packets of the video stream when filtering is activated (as determined by the user-defined rule(s) and the state of the environment 106). For example, the video stream includes, in one embodiment, I-frames and P-frames, and the quality of the video stream is reduced by performing filtering according to any of the following:

- filtering all packets carrying both I-frames and P-frames,
- filtering packets carrying all P-frames (but keeping packets carrying I-frames),

- filtering packets carrying a subset of the P-frames (e.g., but keeping the other P-frames and all I-frames), or
- filtering packets carrying a determined amount of the P-frames (e.g., filtering every Mth P-frame but keeping other P-frames and all I-frames where "M" is determined based on, e.g., the user-defined rule(s) and the state of the environment 106).

[0051] In one embodiment, the system 100 of Figure 1 can be implemented or represented as a modified version of the architecture defined by 3GPP TS 23.434, as illustrated in Figure 2 where new components are represented by bold boxes and new interfaces are represented by bold lines. In other words, Figure 2 (which is an extension of Figure 14.2.2-1 in 3GPP TS 23.434) illustrates an on-network functional model for Network Resource Management (NRM) extended with an NRM action UE aggregation point, which corresponds to the network-controlled device 110 of Figure 1, and an NRM action transport side function.

[0052] As illustrated in Figure 2, the system 100 includes a Vertical Application Layer (VAL) UE 200, a 3GPP network system 202, a VAL server(s) 204, and an NRM server 206. The VAL UE 200 may be implemented within the camera 102-i of Figure 1 (e.g., each of the cameras 102-1 through 102-6 includes its own VAL UE 200). The 3GPP network system 202 corresponds to the 3GPP network 112 of Figure 1, and the VAL server(s) 204 correspond to the cloud server(s) 114 of Figure 1. The NRM server 206 is not illustrated in Figure 1. As further illustrated in Figure 2, the system includes a NRM action UE aggregation point 208, which corresponds to the network-controlled device 110 of Figure 1.

[0053] The VAL UE 200 includes one or more VAL clients 210, a NRM client 212 including a VAL client adaption trigger and router 214, and a NRM action UE 216. The NRM server 206 includes a VAL server adaptation trigger and router 218. The system 100 further includes an NRM action transport 220.

[0054] The new functional entities are: the NRM action UE aggregation point 208, the NRM action UE 216, and the NRM action transport 220. The NRM action UE aggregation point 208 is responsible for executing the NRM action outside of the 3GPP and VAL domains. The NRM action UE aggregation point 208 collects the traffic of a set of VAL UEs 200 and is connected via a radio channel(s) to the 3GPP network system 202.

[0055] The NRM action UE 216 is, in the illustrated example, hosted in the driver, or Operating System (OS), of the VAL UE 200 or as a tunneling node on the network interface of the VAL UE 200.

[0056] The NRM action transport 220 can be hosted on the transport network or the
5 interface between the core network and radio access network of the 3GPP network system 202.

[0057] The interface between the NRM server 206 and the NRM action UE 216 is denoted herein as "NRM-OS". The interface between the NRM client 212 and the NRM action transport 220 is denoted here as "NRM-T". The interface between the NRM client
10 212 and the NRM action UE aggregation point 208 is denoted herein as "NRM-UE Mux". The interface between the NRM server 206 and the NRM action UE aggregation point 208 is denoted herein as "NRM-Transport Mux".

[0058] The following procedures and information flows are performed within the system 100 of Figure 2. Table 1 below shows one example embodiment of an extension
15 of Table 14.3.2.1-1 of 3GPP TS 23.434 for the information flow for a network resource adaptation request from the VAL client 210 to the NRM client 212. The new information elements are highlighted in bold, underlined text. A new information element is the Source IP address and source Port identifier of the video stream. The resource adaptation requirement field should contain video related information. Currently, it is
20 defined as a string in 3GPP TS 29.549 (see, e.g., V18.1.0), thus that field can be utilized without any modification of the standard.

Table 1: Modified version of Table 14.3.2.1-1 of 3GPP TS 23.434 - Client-side Network resource adaptation action request for video stream

Information element	Status	Description
Requester Identity	M	The identity of the VAL server performing the request.
List of VAL UE IDs	O (see NOTE)	List consisting of one or more VAL UE IDs for whom the network resource adaptation occurs.
VAL group ID	O (see NOTE)	The VAL group ID for whom the network resource adaptation occurs.
<u>Source IP, source Port identifiers</u>	<u>M</u>	<u>The source IP, port of the video stream</u>
Resource adaptation requirement	M	The resource adaptation requirement corresponds to the VAL service QoS requirements as applied for a UE or group of UEs (E.g. bandwidth, resource, <u>FPS, resolution</u>).
NOTE: Either of the information elements should be present.		

[0059] Table 2 below shows one example embodiment of an extension of Table 14.3.2.2-1 of 23.434 that describes the information flow for a network resource adaptation response from the NRM client 212 to the VAL client 210. The new information elements are highlighted in bold, underlined text. The response can provide information on a failure (e.g., if applicable) regarding a capability limitation(s) of a certain network element. For example, in case the NRM action UE aggregation point 208 cannot handle the NRM request, then the VAL client adaptation trigger and router 214 can route the request to the NRM action UE 216.

Table 2: Modified version of Table 14.3.2.2-1 of 3GPP TS 23.434 - Client-side Network resource adaptation action response for video stream

Information element	Status	Description
Result	M	Result includes success or failure of the network resource adaptation with the underlying network. The response can also include an updated value for some of the parameters included in the network resource adaptation request (e.g. negotiation of resource offering). <u>The result can also include indication of the failure regarding the capability limitation of a certain network element.</u>

[0060] Regarding a resource adaptation request (referred to herein as "NRM-UE Mux request") from the NRM client 214 to the NRM action UE aggregation point 208, the information elements on this interface can have less information than described above for the network resource adaptation action request for a video stream as we skip out of the 3GPP defined VAL domain. In one embodiment, the action needs the source IP address, source Port identifiers, and the resource adaptation requirement, as illustrated in Table 3 below. The format of the message, and the transport protocol is preferred to be in byte format e.g., protobuf over UDP than e.g., verbose json format, as the in-network computing boxes have limited string parsing capabilities.

Table 3

Information element	Status	Description
<u>Source IP, source Port identifiers</u>	<u>M</u>	<u>The source IP, port of the video stream</u>
<u>Resource adaptation requirement</u>	<u>M</u>	<u>The resource adaptation requirement corresponds to the VAL service QoS requirements as applied for a UE or group of UEs (E.g. bandwidth, resource, FPS, resolution, low/high quality).</u>

[0061] Table 4 below illustrates an example embodiment of an extension of Table 14.3.2.2-1 of 3GPP TS 23.434 that describes the information flow network resource

- 5 adaptation response (referred to herein as a "NRM-UE Mux response") from the NRM action UE aggregation point 208 to the NRM client 214. The result can provide information on the failure regarding the capability limitation of a certain network element. For example, in case the NRM action UE aggregation point 208 cannot handle the NRM request, this information may be provided.

Table 4: Modified version of Table 14.3.2.2-1 3GPP TS 23.434 - Client-side Network resource adaptation action response for video stream

Information element	Status	Description
Result	M	Result includes success or failure of the network resource adaptation with the underlying network. The response can also include an updated value for some of the parameters included in the network resource adaptation request (e.g. negotiation of resource offering). <u>The result can also include indication of the failure regarding the capability limitation of a certain network element.</u>

[0062] As illustrated in Table 5 below, the NRM-Transport Mux request from the NRM server 206 to the NRM action UE aggregation point 208 is the same as the NRM-UE Mux described above, but it does not need to be radio friendly.

Table 5: NRM-Transport Mux request

Information element	Status	Description
<u>Source IP, source Port identifiers</u>	<u>M</u>	<u>The source IP, port of the video stream</u>
<u>Resource adaptation requirement</u>	<u>M</u>	<u>The resource adaptation requirement corresponds to the VAL service QoS requirements as applied for a UE or group of UEs (E.g. bandwidth, resource, FPS, resolution, low/high quality).</u>

[0063] Table 6 below illustrates one example embodiment of an extension of Table 14.3.2.2-1 of 3GPP TS 23.434 that describes the information flow network resource adaptation response (i.e., the NRM-Transport Mux response) from the NRM action transport 220 to the NRM server 206. The result can provide information on the failure regarding the capability limitation of a certain network element. For example, in case the NRM action transport 220 cannot handle the NRM request, this can be indicated by the failure information.

Table 6: Modified version of Table 14.3.2.2-1 of 3GPP TS 23.434 - Client-side Network resource adaptation action response for video stream

Information element	Status	Description
Result	M	Result includes success or failure of the network resource adaptation with the underlying network. The response can also include an updated value for some of the parameters included in the network resource adaptation request (e.g. negotiation of resource offering). The result can also include indication of the failure regarding the capability limitation of a certain network element.

[0064] Figure 3 illustrates the operation of the system 100 of Figure 2 in accordance with one embodiment of the present disclosure. As illustrated, the VAL client 210 sends a resource adaptation request for a video stream to the NRM client 212 (step 300). In one embodiment, the resource adaptation request includes the source IP address and port identifier(s) of the video stream, e.g., as described above in relation to Table 1. The NRM client 212 sends a resource adaptation request to the NRM server 206 (step 304). The resource adaptation request sent from the NRM client 212 to the NRM server 206 is the same as or includes the resource adaptation request received from the VAL client 212 or includes at least some of the information from the resource adaptation

request (e.g., source IP address and port identifier(s) of the video stream) received from the VAL client 212. Based on the resource adaptation request, the NRM server 206 sends an adaptation request for the video stream to the NRM action UE aggregation point 208 (step 308). The adaptation request is a request to activate adaptation for the video stream (e.g., identified by the source IP address and port identifier(s) of the video stream) indicated by received resource adaptation request. As discussed above, the adaptation request (e.g., NRM-Transport Mux request) may also include one or more resource adaptation requirements. In response to the adaptation request, the NRM action UE aggregation point 208 activates adaptation for the indicated video stream (step 310). Note that, in this context, activating adaptation means that the functionality by which the NRM action UE aggregation point 208 decides whether to adapt (e.g., reduce the quality of) the video stream and, if so, adapts (e.g., filters) the video stream is activated.

[0065] The NRM action UE aggregation point 208 then sends a response or "OK" to the NRM server 206 (step 312). This response may be, for example, the NRM-Transport Mux response described above with respect to Table 6. The NRM server 206 sends a resource adaptation response to the NRM client (step 314). This response can be the same as or include information similar to the client-side network resource adaptation action response described above (see Table 2). The NRM client 212 sends a resource adaptation response for the video stream to the VAL client 210 (step 316). This response may be as described above with respect to Table 2.

[0066] Figures 4A through 4C illustrates the operation of the system 100 of Figure 2 in accordance with some embodiments of the present disclosure. Note that optional steps are represented by dashed lines/boxes. Further, while the actions performed are referred to as "steps", these "steps" are not limited to being performed in the order shown in Figures 4A through 4C; rather, the steps may be performed in any desired order and some of the steps may be performed in parallel, unless explicitly stated or otherwise required.

[0067] As shown, signaling to activate video stream adaptation at the aggregation point 208 (step 401). Note that the aggregation point 208 is also referred to herein as an "aggregation node" as it may be implemented as a separate physical node (e.g., hardware device which may execute software stored in memory) in the system 100. The signaling of step 401 may, for example, be the signaling of Figure 3. Note that step 401

is optional. For example, in one alternative embodiment, adaptation is always active for all video streams.

[0068] Cameras 400-1 through 400-N (which correspond to, e.g., the cameras 102-1 through 102-6 in the example of Figure 1) provide (e.g., send or transmit) continuous video streams to the aggregation point 208 via respective VAL UEs 200-1 through 200-N (steps 402-1 through 502-N and steps 404-1 through 504-N). The aggregation point 208 obtains (e.g., receives) one or more adaptation or filtering rules (e.g., via the API 118) and state information for the environment (e.g., environment 106) in which the cameras 400-1 through 400-N are located (step 406). The state information is, in one embodiment, obtained by observing information included in status updates transmitted by device(s) (e.g., robot 108, sensor(s), etc.) in the environment that pass through the aggregation point 208. The aggregation point 208 determines (e.g., based on the rule(s) and state information) that a frame rate of a certain video stream from a certain camera 400-i (where "i" is used here to denote an arbitrary camera from among the cameras 400-1 through 400-N) is to be adapted (e.g., increased or decreased) (step 408). For example, a rule may be defined that states that the frame rate of a video stream from any of the cameras is to be reduced if the state information indicates that the robot 108 is not within the area captured by the respective camera.

[0069] Responsive to determining that the frame rate of the video stream is to be adapted (e.g., reduced in this example), the aggregation point 208 adapts the frame rate of the video stream from a first frame rate to a lower, second frame rate (step 410). This adaptation may include any of the following:

- filtering (i.e., dropping) all packets carrying both I-frames and P-frames,
- filtering packets carrying all P-frames (but keeping packets carrying I-frames),
- filtering packets carrying a subset of the P-frames (e.g., but keeping the other P-frames and all I-frames), or
- filtering packets carrying a determined amount of the P-frames (e.g., filtering every Mth P-frame but keeping other P-frames and all I-frames where "M" is determined based on, e.g., the user-defined rule(s) and the state of the environment 106).

In one example embodiment, this adaptation includes determining an amount of a certain frame type(s) (e.g., an amount of P-frames) to be filtered (e.g., based on a required service level) (step 410A). In one example embodiment, this adaptation

includes filtering at least some (e.g., the certain amount determined in step 410A) of a certain frame type(s) (e.g., P-frames) (step 410B). The aggregation point 208 transmits the adapted video stream to, in this example, the VAL server 204 via the 3GPP network system 202 (step 412).

5 **[0070]** The aggregation point 208 may thereafter obtain (e.g., receive or observe) updated state information (e.g., via packeting monitoring/inspection) and/or updated rule(s) (step 414). The aggregation point 208 determines, based on the updated state information and/or updated rule(s), that adaptation of the frame rate of the certain video stream from camera 400-i is to stop (step 416). In response to the determination
10 of step 416, the aggregation point 508 may, in some embodiments, continue adapting (i.e., filtering) the video stream and transmitting the adapted video stream, until reaching a certain frame (e.g., next I-frame) of the video stream (steps 418 and 420).

[0071] The aggregation point 208 stops adaptation (i.e., stops filtering in this example) of the frame rate of the video stream (step 422). Adaptation is stopped, in
15 one example, upon reaching a certain frame (e.g., the next I-frame) of the video stream. In another example, adaptation is stopped immediately or a certain amount of time after making the determination in step 418. The aggregation point 208 transmits the non-adapted video stream to, in this example, the VAL server 204 via the 3GPP network system 202 (step 424).

20 **[0072]** The aggregation point 208 may continue this process to, e.g., dynamically adapt the video streams from the cameras 400-1 to 400-N, e.g., based on the rule(s) and state information of the environment (step 426).

[0073] Figure 5 is a flow chart that illustrates the operation of the aggregation point 208 to adapt the video streams in accordance with one example embodiment of the
25 present disclosure. Note that this process is performed on a per-packet basis (e.g., per IP packet basis where each IP packet contains at least part of a frame of a particular video stream). As illustrated, the aggregation point 208 receives a packet (step 500) and identifies the video stream (e.g., Stream Identifier (SID) contained in the packet) (step 502). The aggregation point 508 selects a frame rate reduction policy and
30 associated parameter(s) (step 504). The frame rate reduction policy may be the same for all video streams or different for different video streams. For example, the frame rate reduction policy may be any of the following:

- drop all frames (e.g., both I-frames and P-frames),

- drop all frames of a certain type(s) (e.g., drop all P-frames), or
- drop a subset of frames of a certain type(s), e.g.:
 - drop every Mth P-frame where "M" is defined by the associated parameter(s), or
 - 5 ○ drop P-frames with a probability given by the associated parameter(s).

In this example, the selected policy is to drop all or a subset of P-frames.

[0074] The aggregation point 208 determines whether the received packet is the start of a P-frame of the video stream (step 506). As will be understood by those of ordinary skill in the art, this can be done by examining the received packet (e.g.,
 10 examining the "S" bit of the Fragmented Unit (FU) header of a Real Time Protocol (RTP) packet including a FU payload for an H.264 or H.265 video stream). If the received packet is not the start of a P-frame of the video stream (step 506, NO), the process proceeds to step 510. If the received packet is the start of a P-frame of the video stream (step 506, YES), the aggregation point 208 determines whether to drop this P-
 15 frame based on the selected policy and the associated parameters and sets an associated drop flag for the video stream accordingly (e.g., sets register[SID]=0 for 'keep and sets register[SID]=1 for 'drop') (step 508).

[0075] Whether proceeding from step 506 or step 508, the aggregation point 208 determines whether the drop flag is set to 'drop' (e.g., is register[SID]=1) (step 510). If
 20 not, the packet is passed (e.g., kept) and proceeds with normal packet processing to be transmitted by the aggregation point 208. However, if the drop flag is set to 'drop' (step 510, YES), the aggregation point 208 marks the packet to drop such that the packet is dropped, or filtered, by the aggregation point 208 (or alternatively drops the packet) (step 512). The aggregation point 208 determines whether the packet is the end of a P-
 25 frame (step 514). As will be understood by those of ordinary skill in the art, this can be done by examining the received packet (e.g., examining the "E" bit of the FU header of a RTP packet including a FU payload for an H.264 or H.265 video stream). If the packet is not the end of a P-frame (step 514, NO), the packet processing continues. If the packet is the end of a P-frame (step 514, YES), the aggregation point 208 sets the drop
 30 flag to 'keep' (e.g., sets register[SID]=0) (step 516), and then packet processing continues.

[0076] Note that the process of Figure 5 is performed for each received packet and is part of the overall packet processing pipeline of the aggregation point 208.

[0077] Figure 4 is similar to Figure 3 but for an embodiment in which the aggregation point 208 drops all P-frames when low-FPS is required. Again, note that this process is performed on a per-packet basis (e.g., per IP packet basis where each IP packet contains at least part of a frame of a particular video stream). As illustrated, the aggregation point 208 receives a packet (step 500) and identifies the video stream (e.g., SID contained in the packet) (step 502). The aggregation point 508 determines whether low-FPS is required for the video stream (step 504'). For instance, if the video stream is a video stream for which the aggregation point 208 determines that a reduction in quality is needed (e.g., in step 408 of Figure 4A), then the aggregation point 208 determines, in step 504', that a low-FPS is needed for the video stream. If low-FPS is not needed (step 504', NO), the packet is passed on for further processing in the packet processing pipeline of the aggregation point 208 for transmission.

[0078] However, if low-FPS is needed (step 504', YES), the aggregation point 208 determines whether the received packet is the start of a P-frame of the video stream (step 506). If not (step 506, NO), the process proceeds to step 510. If the received packet is the start of a P-frame of the video stream (step 506, YES), the aggregation point 208 sets an associated drop flag for the video stream to 'drop' (e.g., sets register[SID]=1 for 'drop') (step 508').

[0079] Whether proceeding from step 506 or step 508, the aggregation point 208 determines whether the drop flag is set to 'drop' (e.g., is register[SID]=1) (step 510). If not, the packet is passed (e.g., kept) and proceeds with normal packet processing to be transmitted by the aggregation point 208. However, if the drop flag is set to 'drop' (step 510, YES), the aggregation point 208 marks the packet to drop such that the packet is dropped, or filtered, by the aggregation point 208 (or alternatively drops the packet) (step 512). The aggregation point 208 determines whether the packet is the end of a P-frame (step 514). If not, the packet processing continues. If so (step 514, YES), the aggregation point 208 sets the drop flag to 'keep' (e.g., sets register[SID]=0) (step 516), and then packet processing continues.

[0080] Note that the process of Figure 6 is performed for each received packet and is part of the overall packet processing pipeline of the aggregation point 208.

[0081] Figure 7 illustrates the operation of the system 100 of Figure 2 in accordance with another embodiment of the present disclosure. In this embodiment, the NRM action UE 208 is requested to activate (and thus perform) the adaptation of the video

stream from an associated camera. In this embodiment, the aggregation point 208 is not needed; rather, the VAL UE 200 (which includes the VAL client 210, the NRM client 212, and the NRM action UE 216) adapts the video stream output by the associated camera (if needed) and transmits the resulting video stream, e.g., to the VAL server 204 via the 3GPP network system 202.

[0082] As illustrated in Figure 7, the VAL client 210 sends a resource adaptation request for a video stream to the NRM client 212 (step 700). In one embodiment, the resource adaptation request includes the source IP address and port identifier(s) of the video stream, e.g., as described above in relation to Table 1. The NRM client 212 sends a resource adaptation request to the NRM action UE 216 (step 304). The resource adaptation request sent from the NRM client 212 to the NRM action UE 216 may be the same as or include the resource adaptation request received from the VAL client 212 or includes at least some of the information from the resource adaptation request (e.g., source IP address and port identifier(s) of the video stream) received from the VAL client 212. Based on the resource adaptation request, the NRM action UE 216 activates adaptation for the indicated video stream (step 706). Note that, in this context, activating adaptation means that the functionality by which the NRM action UE 216 decides whether to adapt (e.g., reduce the quality of) the video stream and, if so, adapts (e.g., filters) the video stream is activated. Once activated, the adaptation may be performed by the NRM action UE 216 in accordance with Figure 5 or Figure 6.

[0083] The NRM action UE 216 then sends a response or "OK" to the NRM client 212 (step 708). This response may, for example, be similar to the NRM-UE Mux response described above with respect to Table 6. The NRM client 212 sends a resource adaptation response for the video stream to the VAL client 210 (step 710). This response may be as described above with respect to Table 2.

[0084] Figures 8A and 8B illustrate the operation of the system 100 of Figure 2 in accordance with some embodiments of the present disclosure. Note that optional steps are represented by dashed lines/boxes. Further, while the actions performed are referred to as "steps", these "steps" are not limited to being performed in the order shown in Figures 8A 8B; rather, the steps may be performed in any desired order and some of the steps may be performed in parallel, unless explicitly stated or otherwise required.

[0085] As shown, signaling to activate video stream adaptation at the NRM action UE 216 (step 801). The signaling of step 801 may, for example, be the signaling of Figure 7. Note that step 801 is optional. For example, in one alternative embodiment, adaptation is always active.

5 **[0086]** Camera 800 (which correspond to, e.g., one of the cameras 102-1 through 102-6 in the example of Figure 1) provides (e.g., sends or transmits) a continuous video stream to the VAL UE 200 (step 602). Note that the following actions shown in Figures 8A and 8B as being performed by the VAL UE 200 are, more specifically, performed by the NRM action UE 212 of the VAL UE 200. The VAL UE 200 obtains (e.g., receives) one
10 or more adaptation or filtering rules (e.g., via an API similar to the API 118) and state information for the environment (e.g., environment 106) in which the camera 800 is located (step 804). Note that the VAL client 210 or VAL server 204 trigger the adaptation and can have logic to trigger the NRM based on the state information of the environment and the rule(s) (e.g., collected by Programmable Logic Controller (PLC) or
15 Open Platform Communication Unified Architecture (OPC-UA), or any kind of Industrial Internet of Things (IIoT) protocol). The VAL UE 200 determines (e.g., based on the rule(s) and state information) that a frame rate of the video stream from the camera 800 is to be adapted (e.g., increased or decreased) (step 806). For example, a rule may be defined that states that the frame rate of the video stream is to be reduced if the
20 state information indicates that the robot 108 is not within the area captured by the respective camera.

[0087] Responsive to determining that the frame rate of the video stream is to be adapted (e.g., reduced in this example), the VAL UE 200 adapts the frame rate of the video stream from a first frame rate to a lower, second frame rate (step 808). This
25 adaptation may include any of the following:

- filtering (i.e., dropping) all packets carrying both I-frames and P-frames,
- filtering packets carrying all P-frames (but keeping packets carrying I-frames),
- filtering packets carrying a subset of the P-frames (e.g., but keeping the other P-frames and all I-frames), or
- 30 • filtering packets carrying a determined amount of the P-frames (e.g., filtering every Mth P-frame but keeping other P-frames and all I-frames where "M" is determined based on, e.g., the user-defined rule(s) and the state of the environment 106).

In one example embodiment, this adaptation includes determining an amount of a certain frame type(s) (e.g., an amount of P-frames) to be filtered (step 808A). In one example embodiment, this adaptation includes filtering at least some (e.g., the certain amount determined in step 808A) of a certain frame type(s) (e.g., P-frames) (step 808B). The VAL UE 200 transmits the adapted video stream to, in this example, the VAL server 204 via the 3GPP network system 202 (step 810). Note that the adaptation may be performed, in one embodiment, in accordance with the process of Figure 5 or Figure 6.

[0088] The VAL UE 200 may thereafter obtain (e.g., receive or observe) updated state information (e.g., via packeting monitoring/inspection) and/or updated rule(s) (step 812). The VAL UE 200 determines, based on the updated state information and/or updated rule(s), that adaptation of the frame rate of the video stream is to stop (step 814). In response to the determination of step 814, the VAL UE 200 may, in some embodiments, continue adapting (i.e., filtering) the video stream and transmitting the adapted video stream, until reaching a certain frame (e.g., next I-frame) of the video stream (steps 816 and 818).

[0089] The VAL UE 200 stops adaptation (i.e., stops filtering in this example) of the frame rate of the video stream (step 820). Adaptation is stopped, in one example, upon reaching a certain frame (e.g., the next I-frame) of the video stream. In another example, adaptation is stopped immediately or a certain amount of time after making the determination in step 816. The VAL UE 200 transmits the non-adapted video stream to, in this example, the VAL server 204 via the 3GPP network system 202 (step 822).

[0090] The VAL UE 200 may continue this process to, e.g., dynamically adapt the video stream, e.g., based on the rule(s) and state information of the environment.

[0091] It should be noted that Figures 3 and Figures 4A through 4B illustrate embodiments of a procedure in which the aggregation point 208 performs the video stream adaptation, whereas Figures 7 and Figures 8A and 8B illustrate embodiments of a procedure in which the NRM action UE 212 at the VAL UE 200 performs the video stream adaptation. In one embodiment, the NRM client 212 or the NRM server 206 selects the route in which the adaptation request is forwarded and thus whether the video stream adaptation is performed by the aggregation point 208 or the NRM action UE 216. This selection may be based on a manual configuration or automated, e.g.,

based on the capability of the camera(s) or the aggregation point 208. One example of an automated process is as follows:

1. The capabilities of the devices (e.g., the camera(s), the VAL UE 200, and/or the aggregation point 208) are collected. For example, the NRM client 212 or the NRM server 206 may check whether there is an aggregation point 208 available for use video stream adaptation for the camera(s). This may be performed by, for example, first determining whether there is a switch available and then obtaining capability information for the switch by, e.g., querying the switch (e.g., via Simple Network Management Protocol (SNMP)) for its device name and model and checking known capabilities of the switch based on its device name and model.
2. If there is a capable aggregation point 208, then the preferred operation is to handle the resource adaptation request in the aggregation point 208. Note that there is no standardized way to upload a P4 code to an in-network switch and run it securely without any interruption of the current operation. The control plane is usually running on an x86; thus any kind of high-level protocol, e.g., HTTP or COAP can be used to upload the code and request an execute. Currently, SSH login and manual triggering of the code is performed by the network engineer.
3. In case of failure (e.g., no capable aggregation point 208 available), the NRM client 212 forwards the adaptation request to the NRM action UE 216.

[0092] The above rule or priority list is practical considering radio-time, or power consumption. If there is any other KPI that alters this priority, it can be used as well.

[0093] In another example embodiment, the adaptation request itself could indicate the place of execution of video stream adaptation.

[0094] Note that the same procedure (e.g., computer program code) can be executed in the aggregation point 208 and the NRM action UE 216 to perform video stream adaptation (e.g., with the help of techniques like Berkley Packet Filter (eBPF) that compiles the P4 code into a kernel module).

[0095] Some example implementation details are shown below for H265-based Real Time Protocol (RTP) streams. In this example, two different profiles are defined for video streams, namely:

- High-FPS: no packet is dropped.

- Low-FPS: drop every packet belonging to a P-frame.

Some other profiles may additionally or alternatively be supported. These other profiles may include any one or more of the following:

- dropping every 2nd P-frame,
- 5 • keeping every nth P-frame,
- dropping P-frames with a given probability,
- detecting and dropping I-frames is also possible,
- any policy that fits the general pipeline described by Figure 3.

[0096] As an example (see, e.g., Figure 6), the frame rate of a video stream is
 10 reduced by dropping every P-frame. Thanks to the I-frames, pictures of the feed still arrive to the receiver. If frames are split between multiple packets, the aggregation point 208 can still detect the frame boundaries based standard flags marking the beginning and the end of the frame.

[0097] In one embodiment, the aggregation point 208 or NRM action UE 216 is
 15 implemented as a programmable switch that is responsible for some or all of the following tasks:

- Maintains IP camera states (e.g., high/low FPS) by checking rule conditions and setting the profiles accordingly;
- Uses match-action tables to read rule parameters;
- 20 • Actions can switch between IP camera states;
- Registers are used to implement traffic filtering;
- In high FPS mode, all packets of the video stream are forwarded;
- In low FPS mode, all (or a subset of) the P-frames are dropped and I-frames (or a part of the I-frames depending on the I-frame rate) are kept (not dropped);
- 25 • The data plane drops entire frames, checking which packets belong together. Note that the beginning and end of I/P-frames are, in some embodiments, identified. Each frame consists of one or more IP packets. Packets corresponding to the beginning and the end of a frame can be easily identified – all packets in between belong to the same frame of the video stream.

30 **[0098]** The adaptation or filtering rules may consider the time between switching profiles. Factors that may be considered include, but are not limited to, speed of movement (e.g., of the robot 108 or device moving within the environment 106), I-frame rate of the video stream, etc.

[0099] Some example use cases are as follows. For a position-based filtering use case, suppose we have a single robot (R) and a single camera (C). The robot reports its position (x,y). The camera observes the rectangular area A defined by (ax,ay) and (bx,by) points. One could deploy the following rule to get high-FPS when the robot is

5 inside the rectangular area:

if $A.ax < R.x < A.bx$ and $A.ay < R.y < A.by$:

SetQuality(C,high-FPS)

else:

SetQuality(C,low-FPS)

10 If there are multiple robots and cameras, the above rules can be efficiently generated and used. Note that A can be modified during runtime thus allowing reorganizing areas and cameras without restarting the aggregation point 208 (e.g., switch).

Use Case: Robot proximity

[0100] For a robot proximity use case, consider a single arbitrarily defined area A that is observed by camera C. Moreover, consider for this example that there are 10 robots each emitting its position and ID. The video stream provided by the camera C is, in this example, to be filtered (i.e., frame rate reduced), unless there are at least 2 robots in the area A observed by the camera C. The following rule solves the problem using a stateful array **arr** and a counter **c** initialized to 0. In this example, **r** always

20 refers to the robot sending the status message under processing.

if **r** in A:

if $arr[r.id] == 0$:

$arr[r.id] = 1$

$c = c + 1$

25 else:

if $arr[r.id] == 1$:

$arr[r.id] = 0$

$c = c - 1$

if $c \geq 2$:

30 $SetQuality(C,high-FPS)$

else:

$SetQuality(C,low-FPS)$

[0101] It should also be noted that, in the embodiments described above, the adaptation of the video stream(s) is performed for the uplink direction (i.e., as video stream(s) are being transmitted in the uplink direction via the cellular network.

However, the same adaptation procedure may additionally or alternatively be performed
5 for the downlink direction. For example, in a scenario in which multiple video streams are being streamed in the downlink direction to a single device (e.g., single UE), an aggregation point may be located in the core network of the cellular communications system to perform adaptation of the video streams prior to transmission of the video streams to the device via the RAN in the downlink direction. The adaptation procedure
10 is the same as that described above and, as such, the details of the adaptation procedure and the aggregation point described above are equally applicable here to the downlink scenario.

[0102] Figure 9 is a schematic block diagram of a network node 900 according to some embodiments of the present disclosure. Optional features are represented by
15 dashed boxes. The network node 900 may be, for example, a network node that implements the aggregation point 208, a network node that implements the VAL server 204, a network node that implements the NRM server 206, or the like. As illustrated, the network node 900 includes a control system 902 that includes one or more processors 904 (e.g., Central Processing Units (CPUs), Application Specific Integrated Circuits
20 (ASICs), Field Programmable Gate Arrays (FPGAs), and/or the like), memory 906, and a network interface 908. The one or more processors 904 are also referred to herein as processing circuitry. In addition, if the network node 900 is a radio access node (e.g., a network controlled router or switch), the network node 900 may include one or more radio units 910 that each includes one or more transmitters 912 and one or more
25 receivers 914 coupled to one or more antennas 916. The radio units 910 may be referred to or be part of radio interface circuitry. In some embodiments, the radio unit(s) 910 is external to the control system 902 and connected to the control system 902 via, e.g., a wired connection (e.g., an optical cable). However, in some other embodiments, the radio unit(s) 910 and potentially the antenna(s) 916 are integrated
30 together with the control system 902. The one or more processors 904 operate to provide one or more functions of the network node 900 as described herein (e.g., one or more functions of the aggregation point 208, the VAL server 204, or the NRM server 206, as described herein). In some embodiments, the function(s) are implemented in

software that is stored, e.g., in the memory 906 and executed by the one or more processors 904.

[0103] In some embodiments, a computer program including instructions which, when executed by at least one processor, causes the at least one processor to carry out the functionality of the network node 900 or a node implementing one or more of the functions of the network node 900 in a virtual environment according to any of the embodiments described herein is provided. In some embodiments, a carrier comprising the aforementioned computer program product is provided. The carrier is one of an electronic signal, an optical signal, a radio signal, or a computer readable storage medium (e.g., a non-transitory computer readable medium such as memory).

[0104] Figure 10 is a schematic block diagram of the network node 900 according to some other embodiments of the present disclosure. The network node 900 includes one or more modules 1000, each of which is implemented in software. The module(s) 1000 provide the functionality of the network node 900 described herein (e.g., one or more functions of the aggregation point 208, the VAL server 204, or the NRM server 206, as described herein).

[0105] Figure 11 is a schematic block diagram of a UE 1100 (e.g., the VAL UE 200) according to some embodiments of the present disclosure. As illustrated, the UE 1100 includes one or more processors 1102 (e.g., CPUs, ASICs, FPGAs, and/or the like), memory 1104, and one or more transceivers 1106 each including one or more transmitters 1108 and one or more receivers 1110 coupled to one or more antennas 1112. The transceiver(s) 1106 includes radio-front end circuitry connected to the antenna(s) 1112 that is configured to condition signals communicated between the antenna(s) 1112 and the processor(s) 1102, as will be appreciated by one of ordinary skill in the art. The processors 1102 are also referred to herein as processing circuitry. The transceivers 1106 are also referred to herein as radio circuitry. In some embodiments, the functionality of the UE 1100 (e.g., the functionality of the VAL UE 200) described above may be fully or partially implemented in software that is, e.g., stored in the memory 1104 and executed by the processor(s) 1102. Note that the UE 1100 may include additional components not illustrated in Figure 11 such as, e.g., one or more user interface components (e.g., an input/output interface including a display, buttons, a touch screen, a microphone, a speaker(s), and/or the like and/or any other components for allowing input of information into the UE 1100 and/or allowing output

of information from the UE 1100), a power supply (e.g., a battery and associated power circuitry), etc.

[0106] In some embodiments, a computer program including instructions which, when executed by at least one processor, causes the at least one processor to carry out the functionality of the UE 1100 according to any of the embodiments described herein is provided. In some embodiments, a carrier comprising the aforementioned computer program product is provided. The carrier is one of an electronic signal, an optical signal, a radio signal, or a computer readable storage medium (e.g., a non-transitory computer readable medium such as memory).

[0107] Figure 12 is a schematic block diagram of the UE 1100 according to some other embodiments of the present disclosure. The UE 1100 includes one or more modules 1200, each of which is implemented in software. The module(s) 1200 provide the functionality of the UE 1100 described herein.

[0108] Any appropriate steps, methods, features, functions, or benefits disclosed herein may be performed through one or more functional units or modules of one or more virtual apparatuses. Each virtual apparatus may comprise a number of these functional units. These functional units may be implemented via processing circuitry, which may include one or more microprocessor or microcontrollers, as well as other digital hardware, which may include Digital Signal Processors (DSPs), special-purpose digital logic, and the like. The processing circuitry may be configured to execute program code stored in memory, which may include one or several types of memory such as Read Only Memory (ROM), Random Access Memory (RAM), cache memory, flash memory devices, optical storage devices, etc. Program code stored in memory includes program instructions for executing one or more telecommunications and/or data communications protocols as well as instructions for carrying out one or more of the techniques described herein. In some implementations, the processing circuitry may be used to cause the respective functional unit to perform corresponding functions according to one or more embodiments of the present disclosure.

[0109] While processes in the figures may show a particular order of operations performed by certain embodiments of the present disclosure, it should be understood that such order is exemplary (e.g., alternative embodiments may perform the operations in a different order, combine certain operations, overlap certain operations, etc.).

[0110] Those skilled in the art will recognize improvements and modifications to the embodiments of the present disclosure. All such improvements and modifications are considered within the scope of the concepts disclosed herein.

Claims

1. A method of operation of radio node for in-network adaptive quality control of video streams from streaming video cameras, the method comprising:
 - for a streaming video camera from among one or more streaming video cameras associated to the radio node:
 - deciding (408; 806) to adapt a frame rate of a video stream received from the streaming video camera;
 - in response to deciding to adapt the frame rate of the video stream received from the streaming video camera:
 - receiving (404-1; 802) the video stream from the stream video camera, the video stream having a first frame rate;
 - adapting (410; 808) a frame rate of the video stream to provide an adapted video stream having a second frame rate that is lower than the first frame rate; and
 - transmitting (412; 810) the adapted video stream at the second frame rate via a wireless network.
2. The method of claim 1 wherein the one or more streaming video cameras associated to the device comprise a plurality of streaming video cameras, the radio node is a network node of the wireless network, and the radio node operates as an aggregation point for video streams from the plurality of streaming video cameras.
3. The method of claim 2 wherein the network node is a network switch.
4. The method of claim 1 wherein the radio node is a wireless communication device enabled to transmit and receive wireless signals to and from the wireless network.
5. The method of any of claim 1 to 4 wherein:
 - the wireless network is a Radio Access Network, RAN, of a cellular communications system;
 - the radio node is a radio node of the cellular communications system;

the method further comprises receiving (308; 704) a request to perform video stream adaptation from a Network Resource Management, NRM, server or an NRM client associated to the cellular communications system; and

5 deciding (408; 806) to adapt the frame rate of the video stream received from the streaming video camera is responsive to receiving (308; 704) the request.

6. The method of claim 5 wherein the radio node is a network node associated to a Radio Access Network, RAN, of the cellular communications system.

10 7. The method of claim 6 wherein the one or more streaming video cameras associated to the device comprise a plurality of streaming video cameras, and the radio node operates as a network-controlled aggregation point for video streams from the plurality of streaming video cameras.

15 8. The method of claim 7 wherein the network node is a network switch.

9. The method of claim 5 wherein the radio node is a User Equipment, UE, enabled to transmit and receive wireless signals to and from a Radio Access Network, RAN, of the cellular communications system.

20

10. The method of any of claims 1 to 9 wherein the video stream comprises both frames of a first frame type and frames of a second frame type that are dependent, directly or indirectly, on the frames of the first frame type, wherein adapting (410; 808) the frame rate of the video stream comprises dropping (410B; 808B) at least some of
25 the frames of the second frame type.

11. The method of claim 10 wherein the frames of the first frame type are key frames, and the frames of the second frame type are predicted frames.

30 12. The method of claim 10 or 11 wherein the video stream is an Internet Protocol, IP, stream comprising a plurality of IP packets, and dropping (410B; 808B) at least some of the frames of the second frame type comprises dropping (410B; 808B) a

subset of the plurality of packets that carry the at least some of the frames of the second frame type.

13. The method of any of claims 10 to 12 wherein dropping (410B; 808B) at least
5 some of the frames of the second frame type comprises dropping (410B; 808B) all frames of the second frame type.

14. The method of any of claims 10 to 12 wherein dropping (410B; 808B) at least
10 some of the frames of the second frame type comprises dropping (410B; 808B) only some frames of the second frame type.

15. The method of any of claims 10 to 12 further comprising:
determining (410A; 808A) an amount of the frames of the second frame type to be dropped;
15 wherein dropping (410B; 808B) at least some of the frames of the second frame type comprises dropping (410B; 808B) the determined amount of the frames of the second frame type.

16. The method of any of claims 1 to 9 wherein the video stream comprises both
20 frames of a first frame type and frames of a second frame type that are dependent, directly or indirectly, on the frames of the first frame type, wherein adapting (410; 808) the frame rate of the video stream comprises dropping both all frames of the first frame type and all frames of the second frame type.

25 17. The method of claim 16 wherein the frames of the first frame type are key frames, and the frames of the second frame type are predicted frames.

18. The method of any of claims 1 to 17 wherein deciding (408; 808B) to adapt the frame rate of the video stream received from the streaming video camera comprises
30 deciding (408; 808B) to adapt the frame rate of the video stream received from the streaming video camera based on one or more predefined rules and state information of one or more secondary devices located with an environment captured by the one or more streaming video cameras.

19. The method of claim 18 wherein the one or more secondary devices move within the environment captured by the one or more streaming video cameras.
- 5 20. The method of claim 19 wherein the one or more secondary devices are robotic devices that move within the environment captured by the one or more streaming video cameras.
- 10 21. The method of claim 19 or 20 wherein the state information of the one or more secondary devices comprises information that indicates locations of the one or more secondary devices within the environment.
- 15 22. The method of claim 18 wherein the one or more secondary devices comprise one or more sensors within the environment captured by the one or more stream video cameras.
- 20 23. The method of any of claims 18 to 22 further comprising obtaining (406; 804) the state information of the one or more secondary devices by monitoring packets transmitted from the one or more secondary devices to an application server via the radio node.
- 25 24. The method of any of claims 1 to 23 further comprising repeating the method for one or more additional stream video cameras from among the one or more streaming video cameras associated to the radio node.
- 30 25. The method of any of claims 1 to 23 further comprising:
deciding (416; 814) to deactivate adaptation of the frame rate of the video stream received from the streaming video camera; and
in response to deciding (416; 814) to deactivate adaptation of the frame rate of the video stream received from the streaming video camera, stopping (422; 820) the adapting of the frame rate of the video stream starting at a certain frame in the video stream.

26. The method of claim 25 wherein the certain frame is a first key frame of the video stream that occurs after deciding to deactivate adaptation of the frame rate of the video stream.

- 5 27. A radio node for in-network adaptive quality control of streaming video cameras, the radio node comprising:
- processing circuitry; and
 - memory comprising instructions executable by the processing circuitry
- 10 whereby the radio node is caused to, for a streaming video camera from among one or more streaming video cameras associated to the radio node:
- decide (408; 806) to adapt a frame rate of a video stream received from the streaming video camera;
 - in response to deciding to adapt the frame rate of the video stream received from the streaming video camera:
- 15 ▪ receive (404-1; 802) the video stream from the stream video camera, the video stream having a first frame rate;
- adapt (410; 808) a frame rate of the video stream to provide an adapted video stream having a second frame rate that is lower than the first frame rate; and
- 20 ▪ transmit (412; 810) the adapted video stream at the second frame rate via a wireless network.

28. A computer program comprising instructions which, when executed on at least one processor, cause the processor to carry out the method according to any of claims
- 25 1 to 26.

29. A carrier containing the computer program of claim 28, wherein the carrier is one of an electronic signal, an optical signal, a radio signal, or a computer readable storage medium.

30

30. A non-transitory computer-readable medium comprising instructions executable by processing circuitry of a radio node whereby the radio node is operable to:

- for a streaming video camera from among one or more streaming video cameras associated to the radio node:
 - decide (408; 806) to adapt a frame rate of a video stream received from the streaming video camera;
 - 5 ○ in response to deciding to adapt the frame rate of the video stream received from the streaming video camera:
 - receive (404-1; 802) the video stream from the stream video camera, the video stream having a first frame rate;
 - adapt (410; 808) a frame rate of the video stream to provide an
10 adapted video stream having a second frame rate that is lower than the first frame rate; and
 - transmit (412; 810) the adapted video stream at the second frame rate via a wireless network.

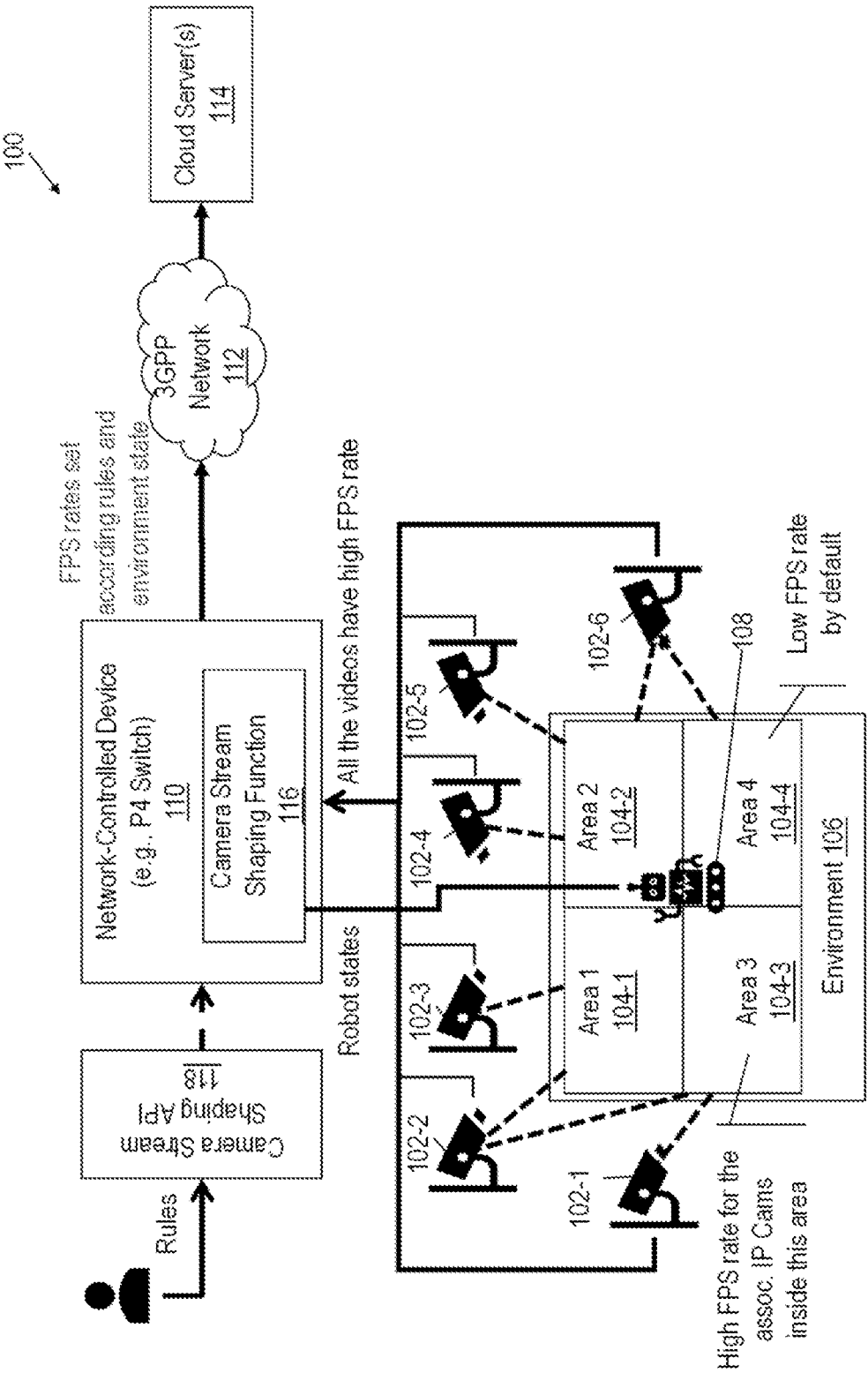


FIG. 1

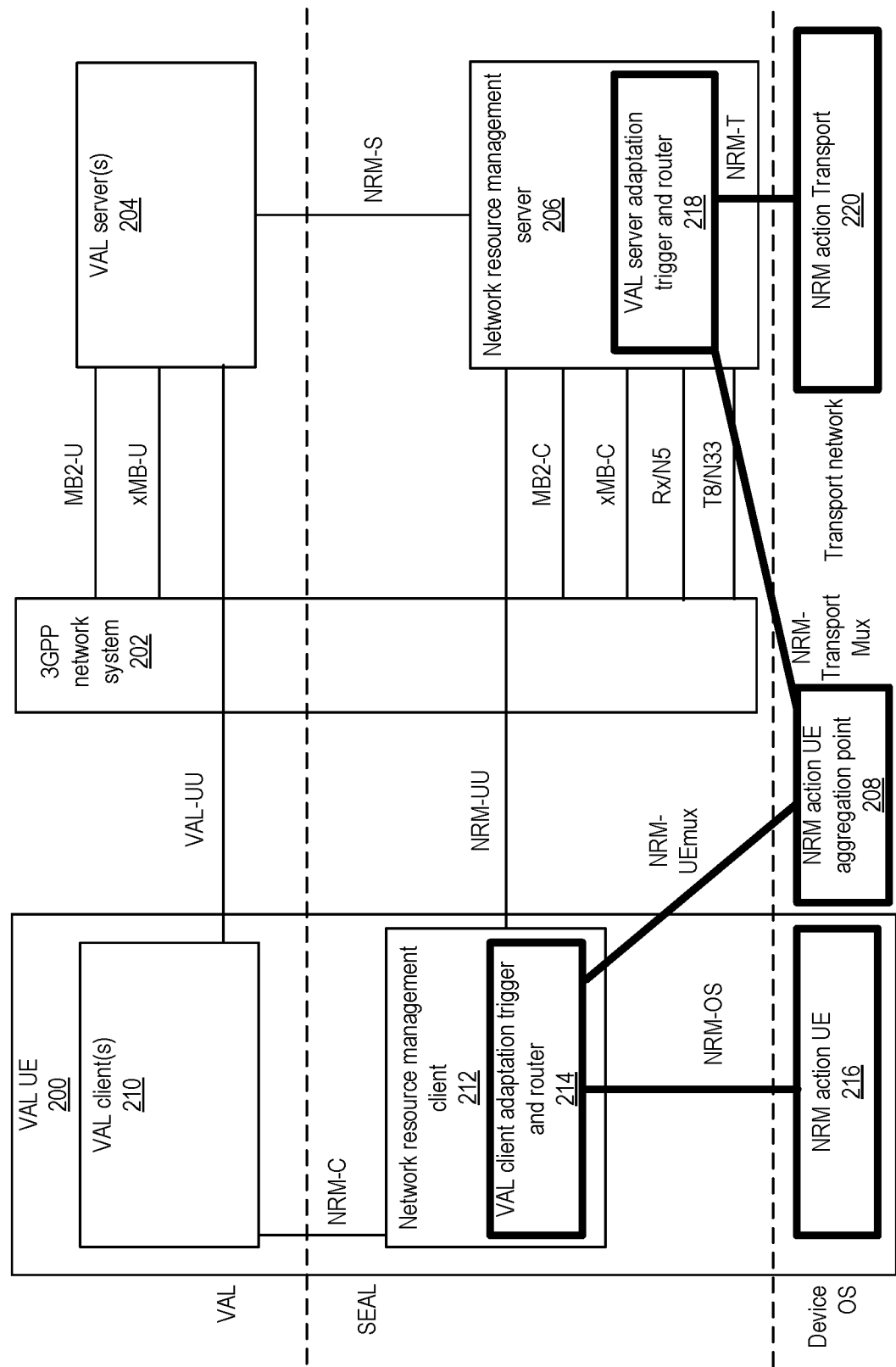


FIG. 2

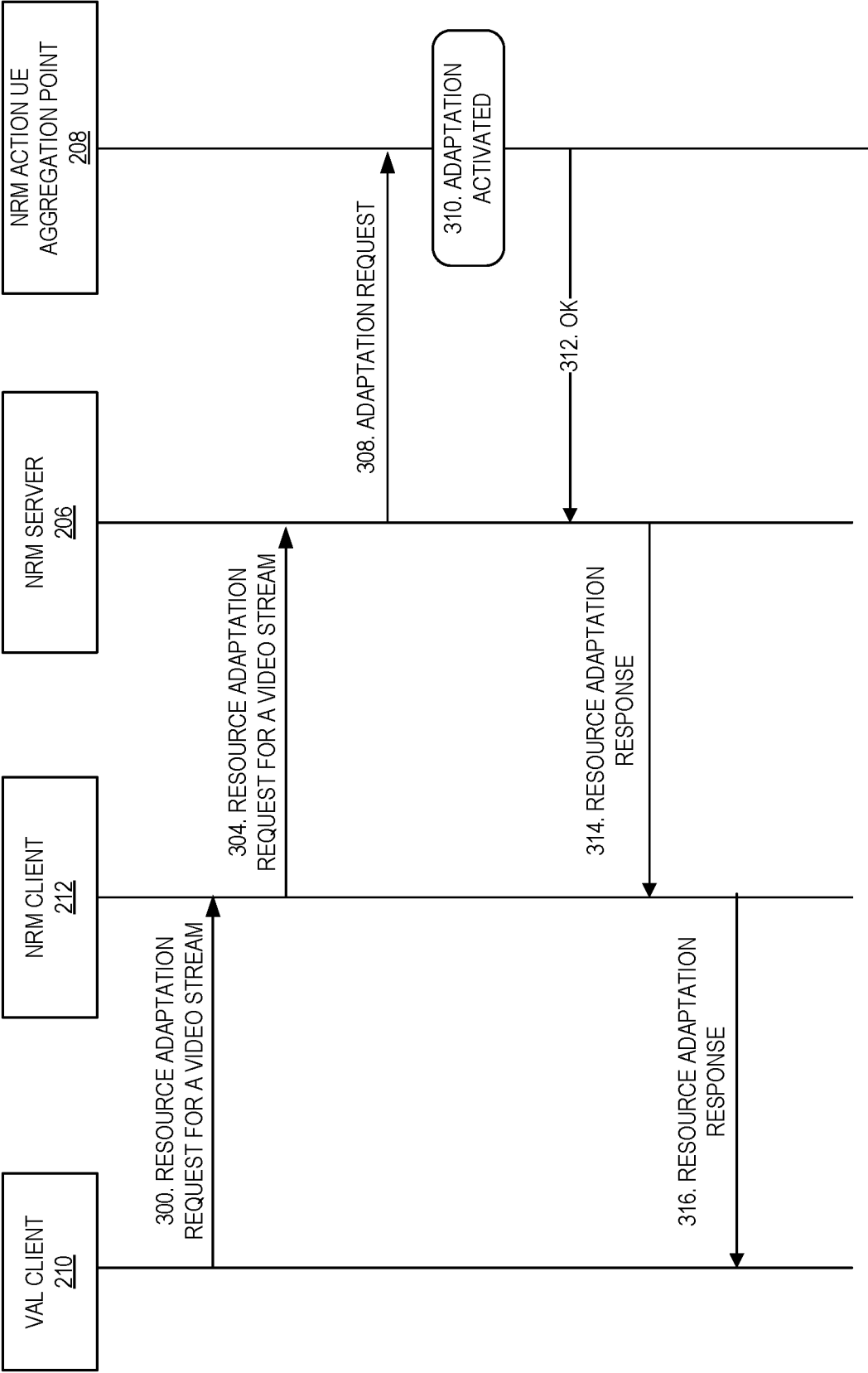


FIG. 3

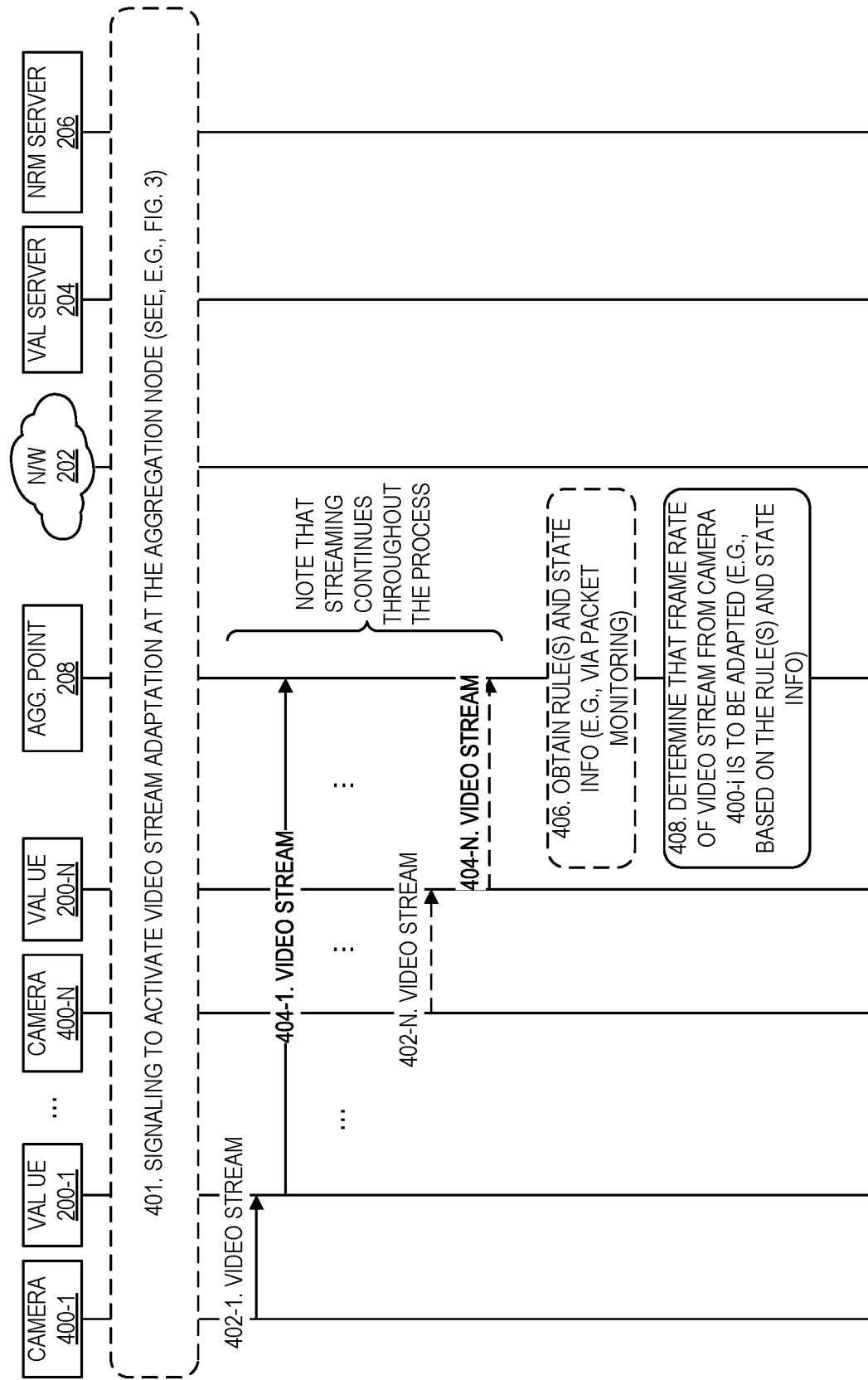


FIG. 4A

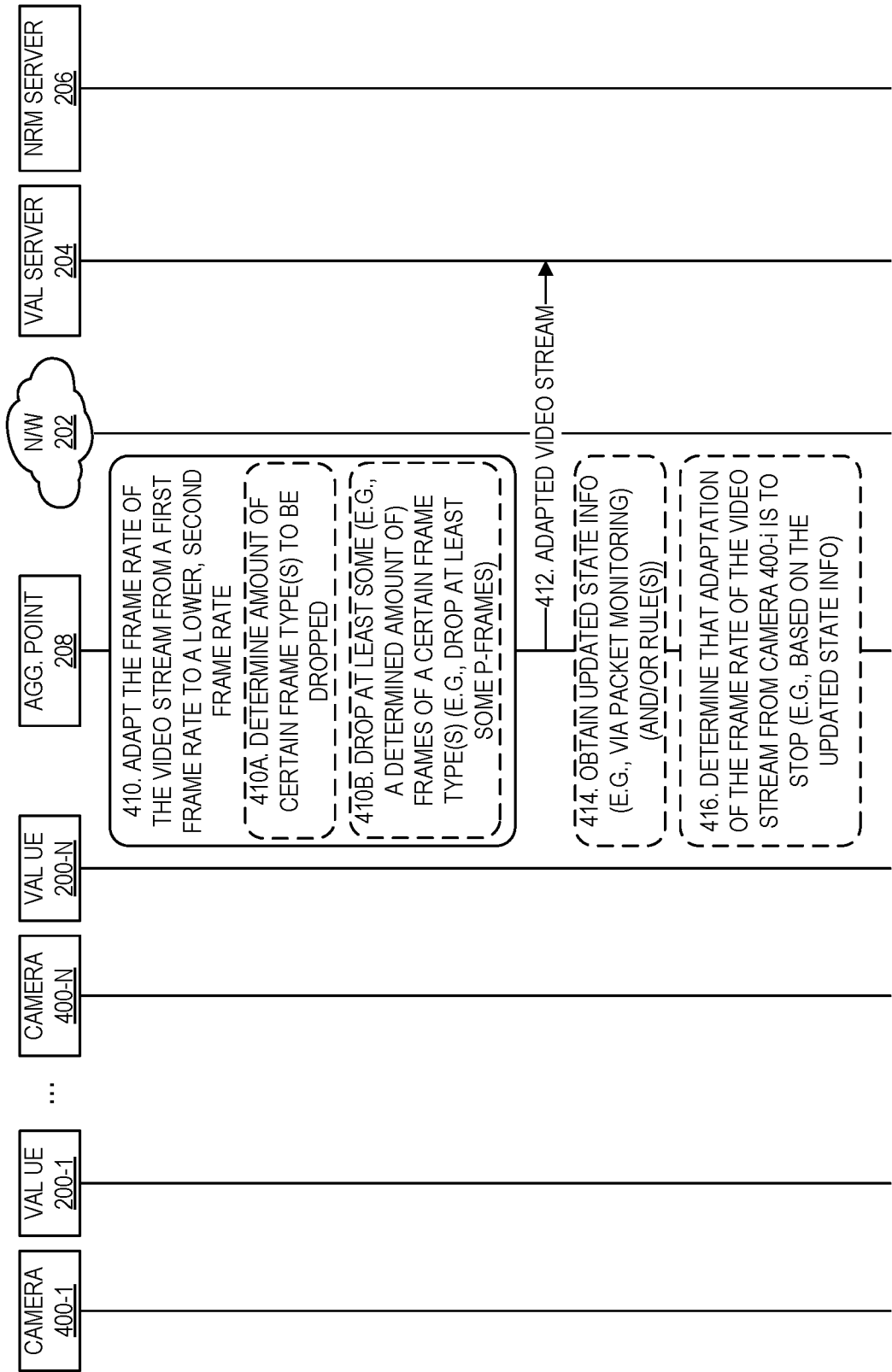


FIG. 4B

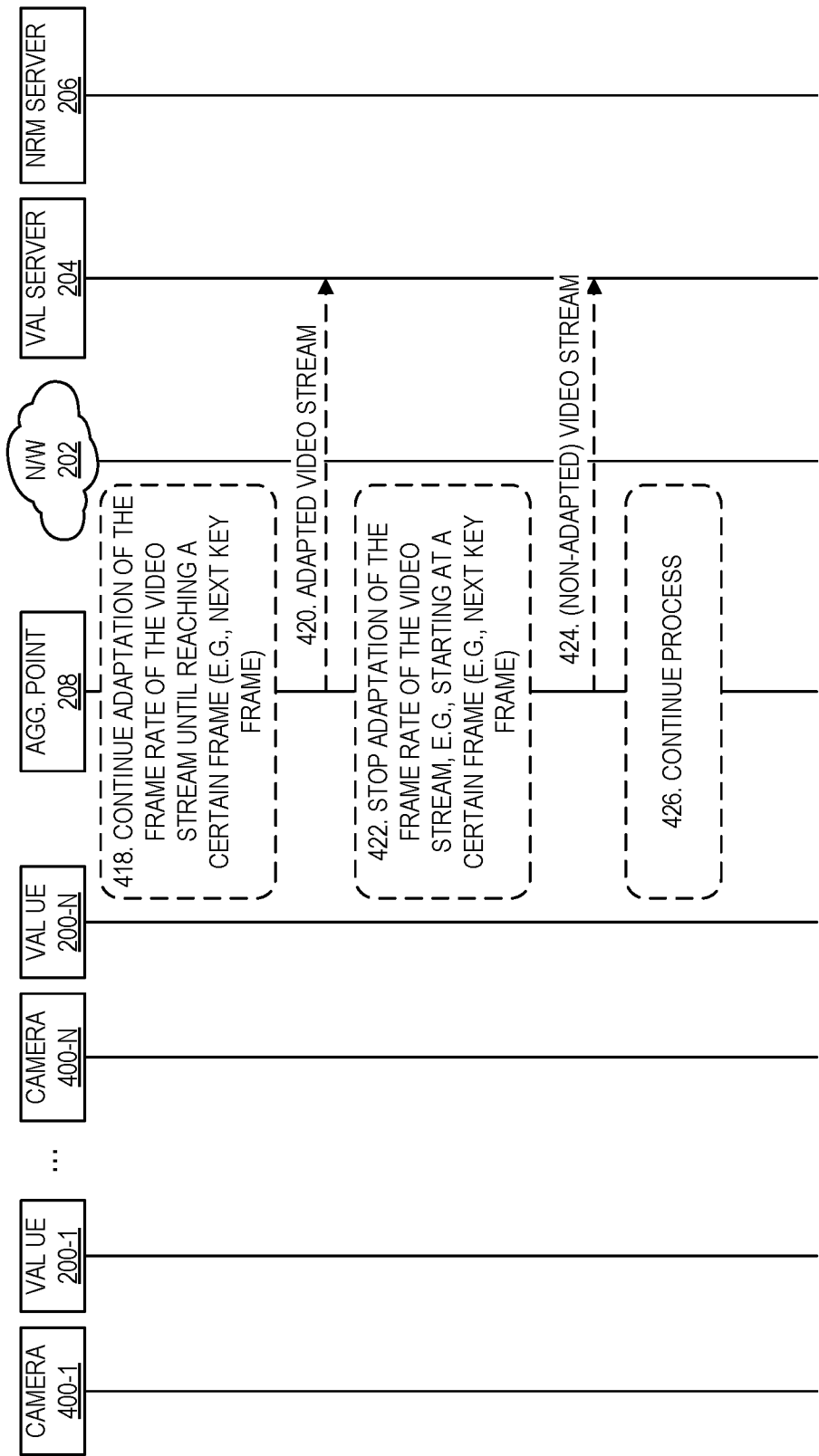


FIG. 4C

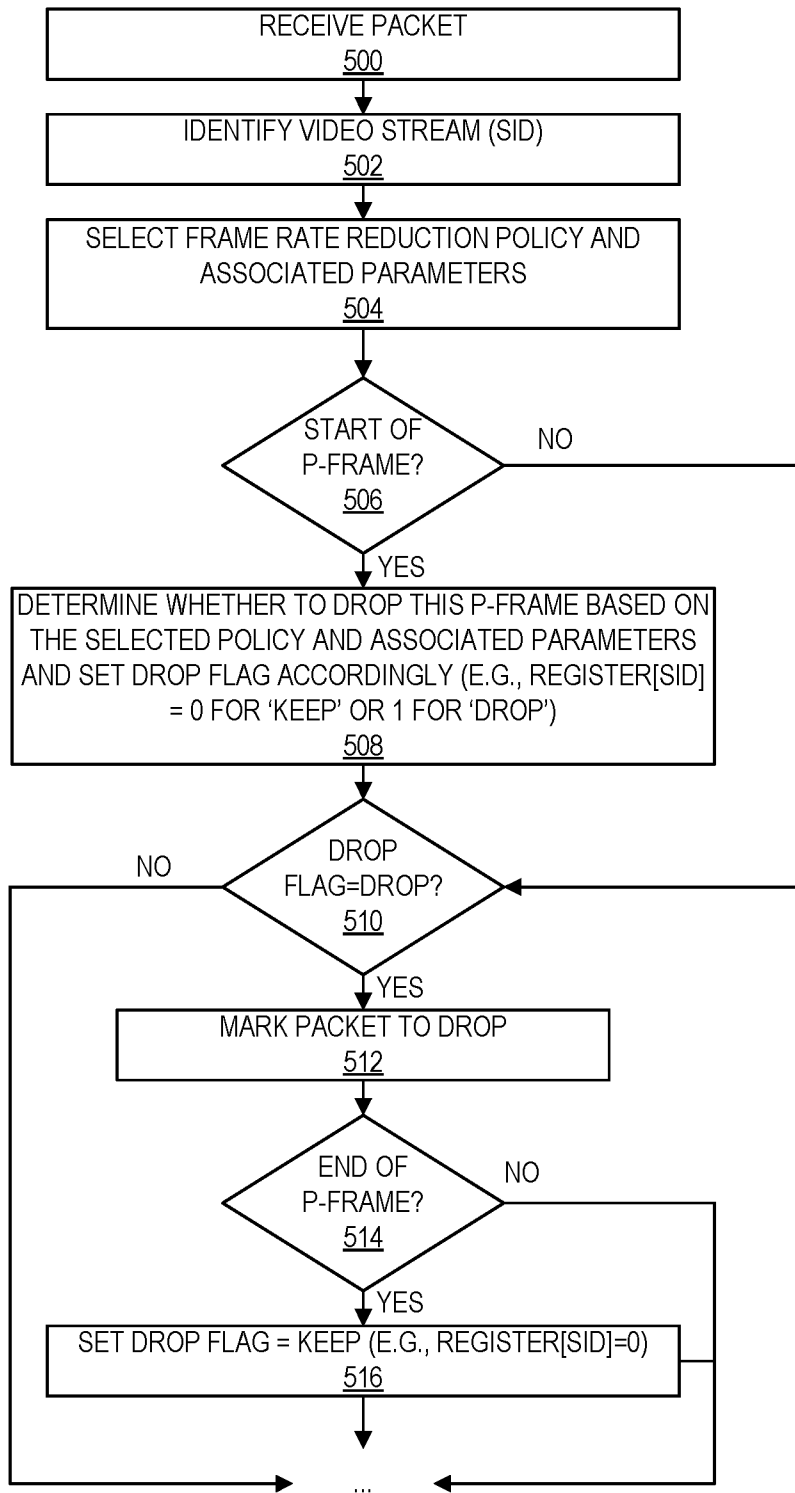
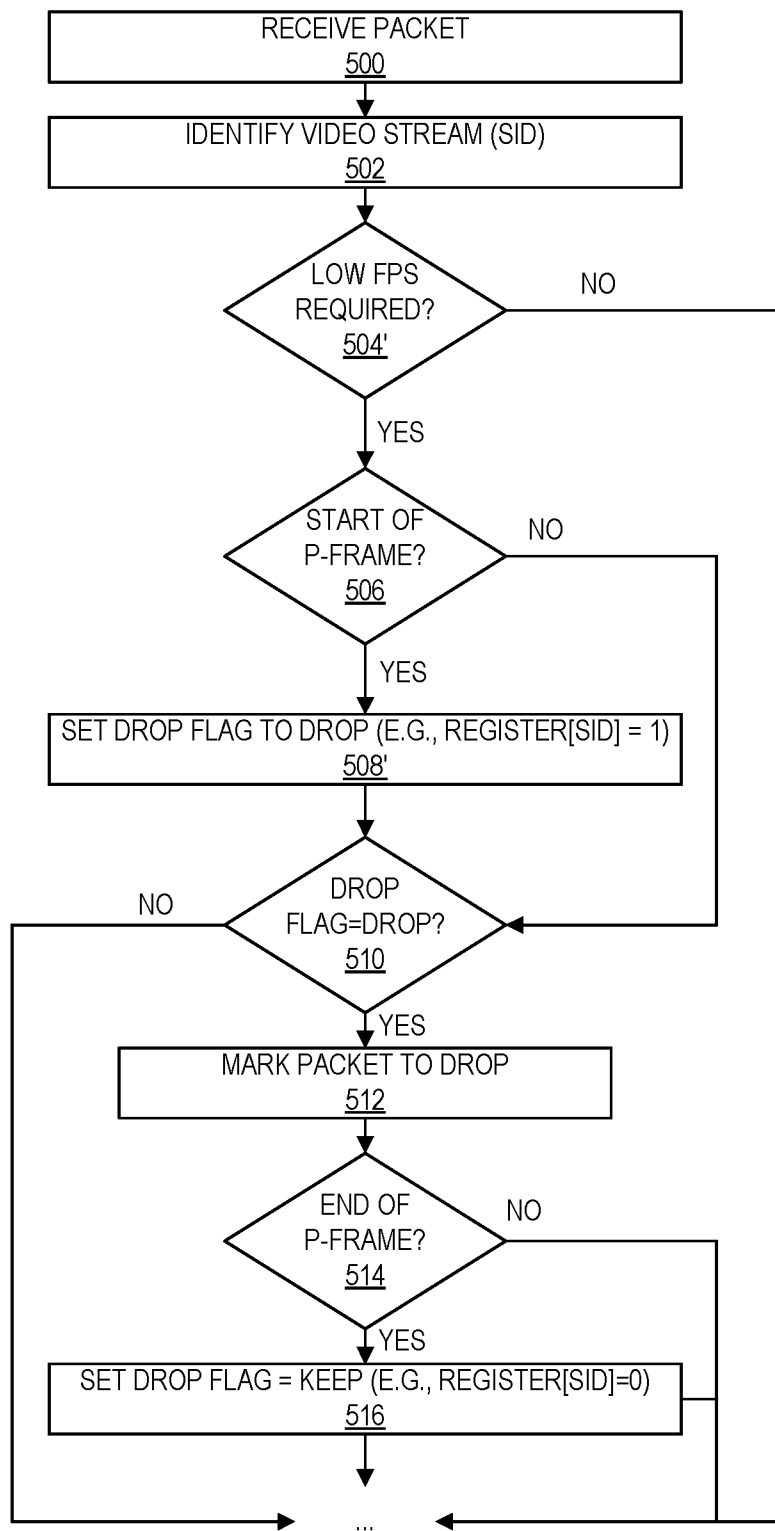


FIG. 5

**FIG. 6**

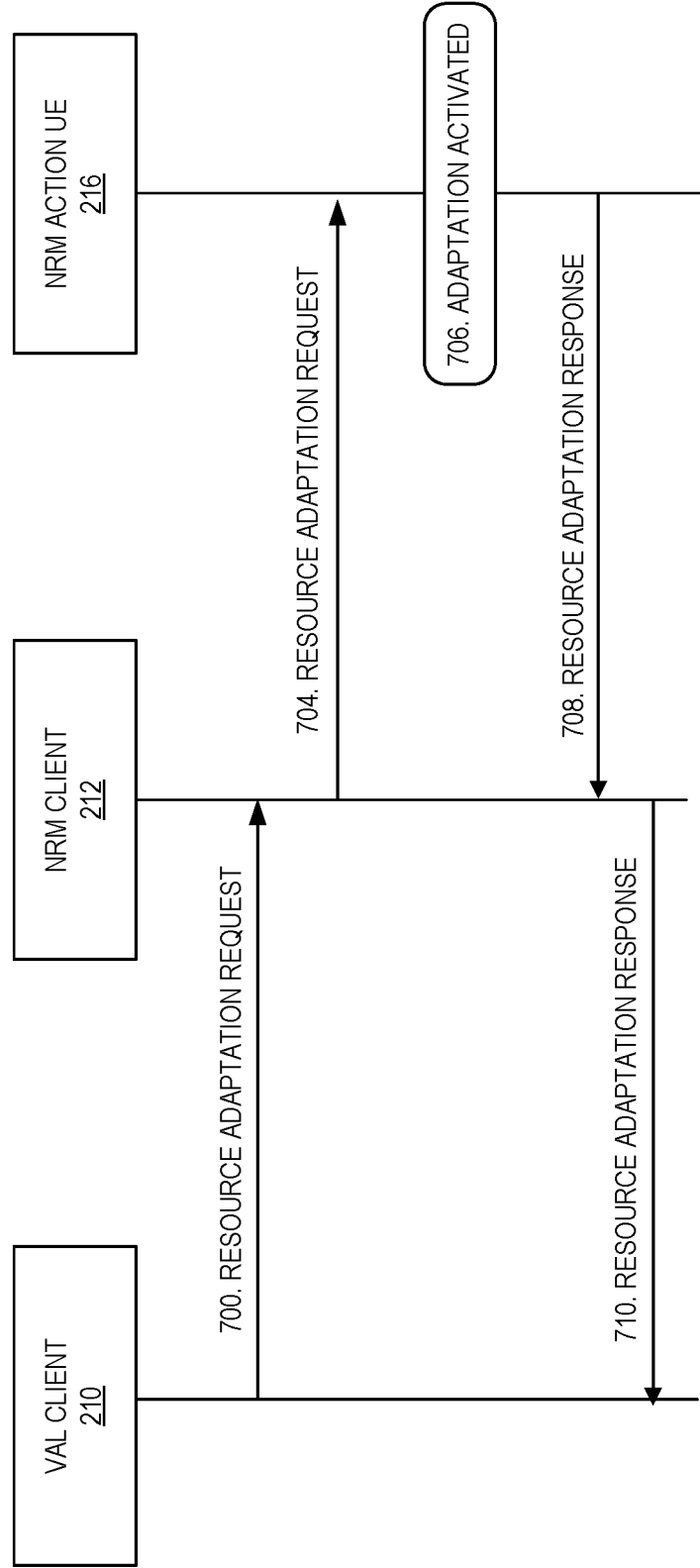


FIG. 7

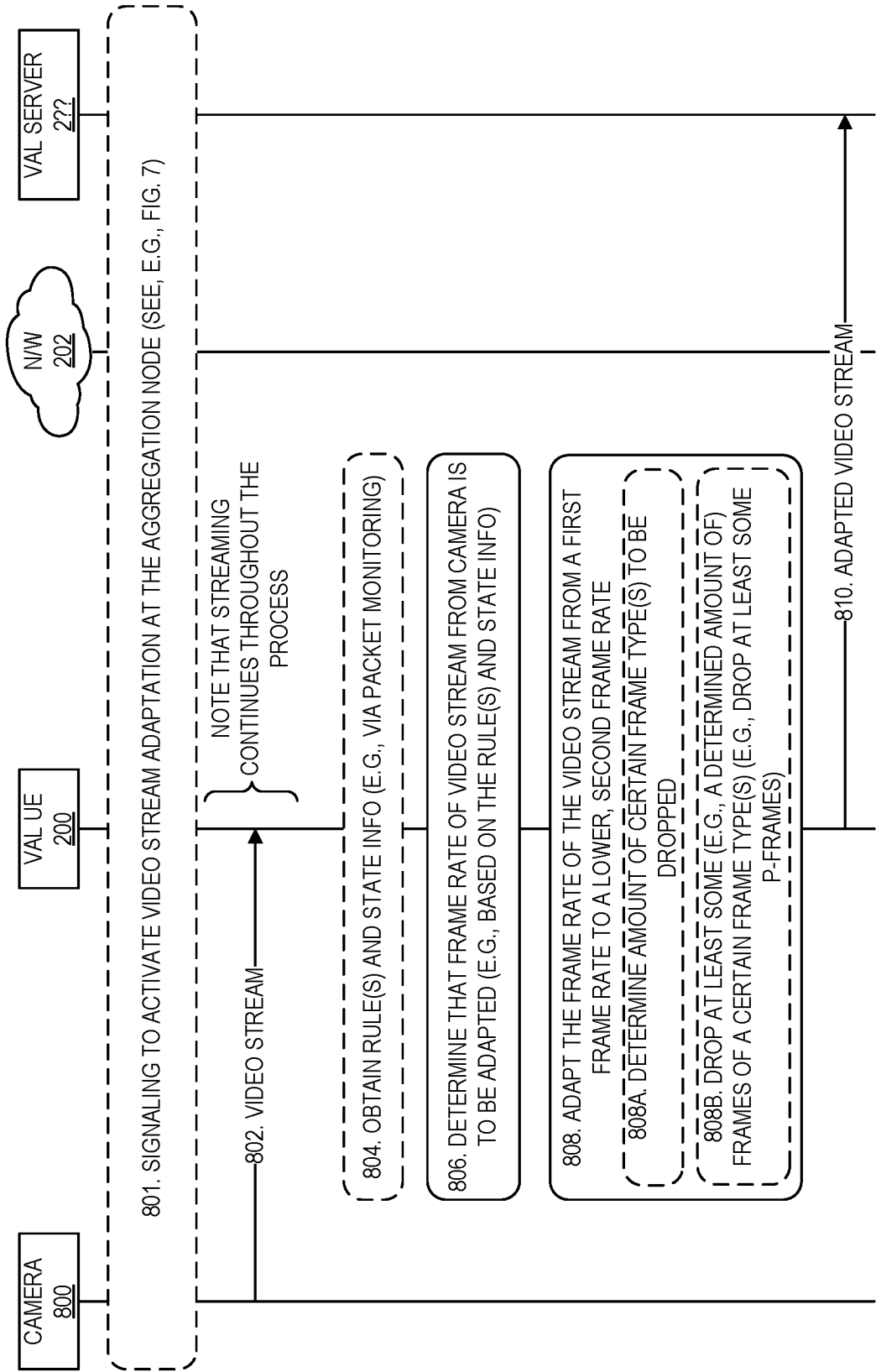


FIG. 8A

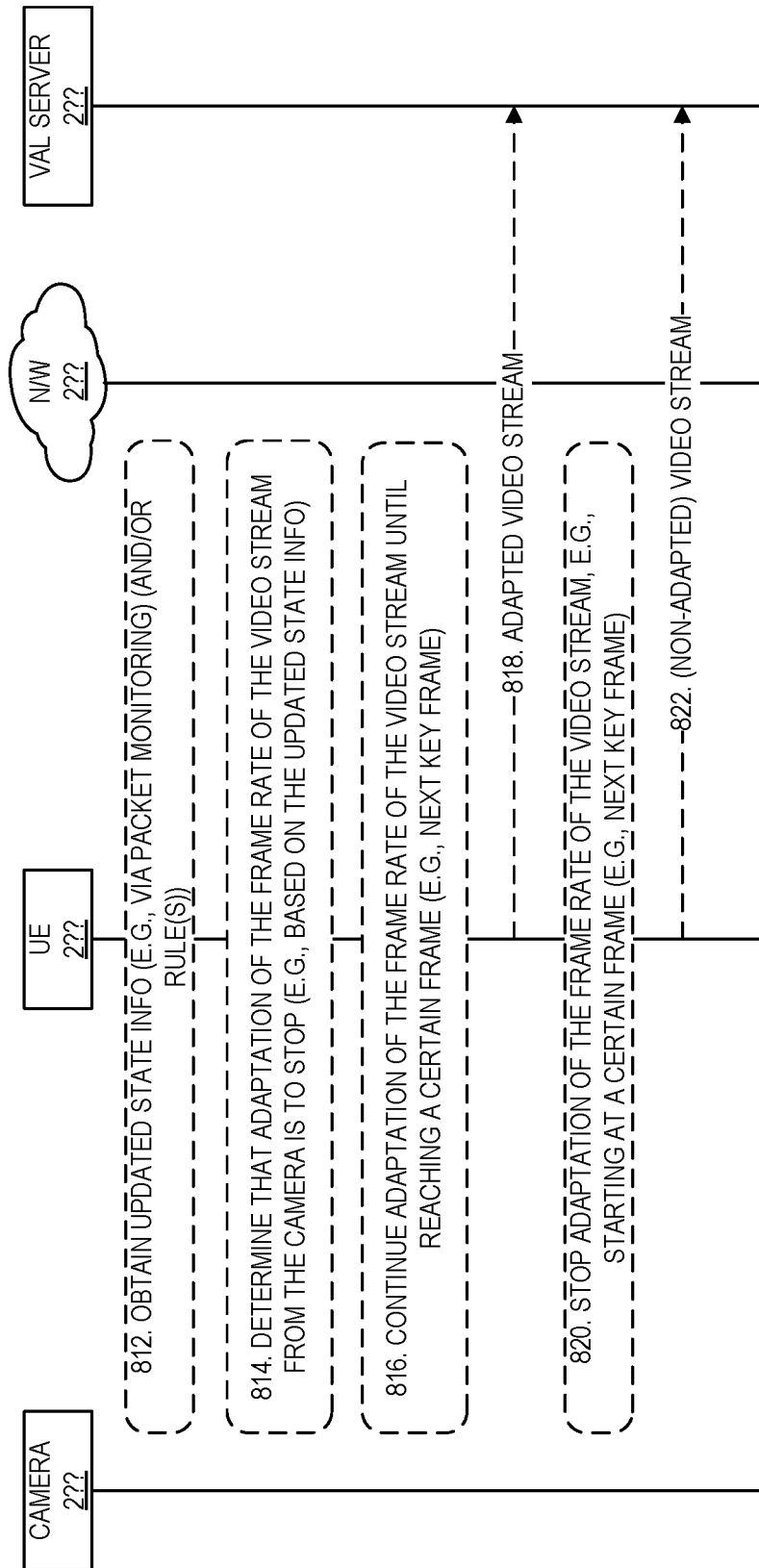


FIG. 8B

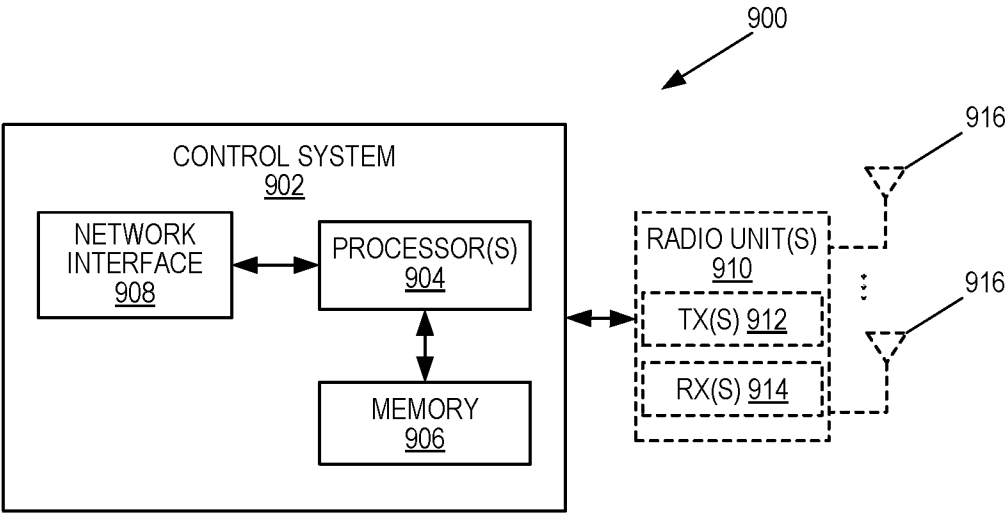


FIG. 9

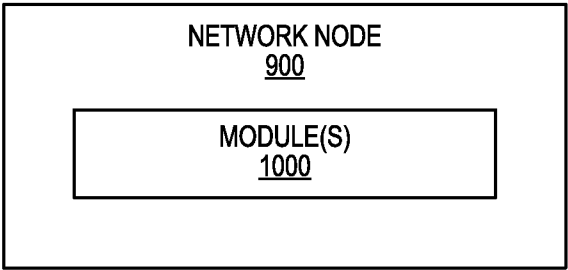
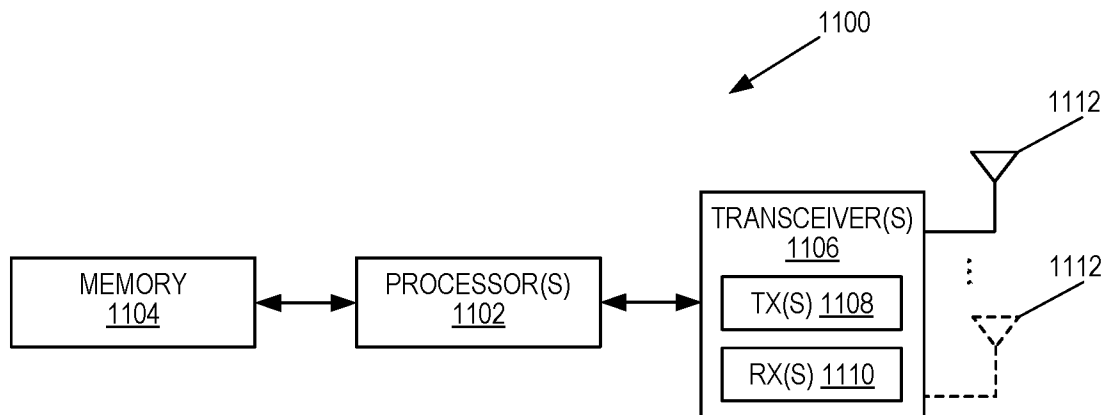
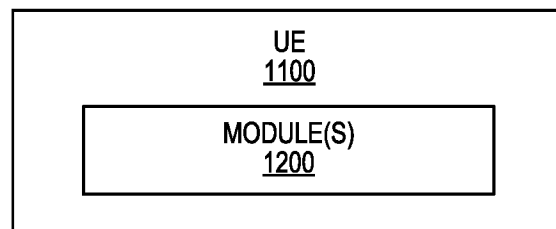


FIG. 10

**FIG. 11****FIG. 12**

INTERNATIONAL SEARCH REPORT

International application No

PCT/IB2023/055782

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N7/18 H04N21/2343 H04N21/647

ADD.

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)

H04N H04L G08B

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)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2019/014388 A1 (RUTLEDGE CHRISTOPHER LEE [US] ET AL) 10 January 2019 (2019-01-10) abstract figures 1,2,4 paragraphs [0018] - [0022], [0029], [0032], [0034], [0038] -----	1-30
A	US 8 776 161 B2 (GAZIER MICHAEL [CA]; DUNCAN IAN [CA] ET AL.) 8 July 2014 (2014-07-08) abstract figures 2,3,4,6 column 5, line 62 - column 6, line 7 column 7, line 25 - column 8, line 37 ----- -/--	1-17, 24-30



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 but 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

5 December 2023

Date of mailing of the international search report

14/12/2023

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2

NL - 2280 HV Rijswijk

Tel. (+31-70) 340-2040,

Fax: (+31-70) 340-3016

Authorized officer

Sery, Franck

INTERNATIONAL SEARCH REPORT

International application No

PCT/IB2023/055782

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2021/281771 A1 (YANG HAI [CN]) 9 September 2021 (2021-09-09) abstract figure 6 paragraphs [0075] - [0080] -----	10, 11
A	US 2017/364755 A1 (WU KE [CN] ET AL) 21 December 2017 (2017-12-21) abstract -----	18-21
A	US 2022/392081 A1 (RYAN DAN [US] ET AL) 8 December 2022 (2022-12-08) abstract paragraph [0017] -----	18-22

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IB2023/055782

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2019014388 A1	10-01-2019	NONE	
US 8776161 B2	08-07-2014	NONE	
US 2021281771 A1	09-09-2021	CN 109379625 A EP 3886448 A1 US 2021281771 A1 WO 2020108082 A1	22-02-2019 29-09-2021 09-09-2021 04-06-2020
US 2017364755 A1	21-12-2017	CN 105975633 A EP 3261019 A1 US 2017364755 A1 WO 2017219506 A1	28-09-2016 27-12-2017 21-12-2017 28-12-2017
US 2022392081 A1	08-12-2022	NONE	