Cloud-based Video Monitoring Framework: An Approach based on Software-Defined Networking for Addressing Scalability Problems

Nay Myo Sandar[1], Sivadon Chaisiri[1], Sira Yongchareon[2]
, and Veronica Liesaputra[2]

[1] School of Information Technology, Shinawatra University, Thailand
55401002-5@st.siu.ac.th, sivadon@ieee.org
[2] Department of Computing, Unitec Institute of Technology, Auckland, New Zealand
sira@maxsira.com, vliesaputra@unitec.ac.nz

Abstract. Closed-circuit television (CCTV) and Internet protocol (IP) cameras have been applied to a surveillance or monitoring system, from which users can remotely monitor video streams. The system has been employed for many applications such as home surveillance, traffic monitoring, and crime prevention. Currently, cloud computing has been integrated with the video monitoring system for achieving value-added services such as video adjustment, encoding, image/video recognition, and backup services. One of the challenges in this integration is due to the size and geographical scalability problems when video streams are transferred to and retrieved from the cloud services by numerous cameras and users, respectively. Unreliable network connectivity is a major factor that causes the problems. To deal with the scalability problems, this paper proposes a framework designed for a cloud-based video monitoring (CVM) system. In particular, this framework applies two major approaches, namely stream aggregation (SA) and software-defined networking (SDN). The SA approach can reduce the network latency between cameras and cloud services. The SDN approach can achieve the adaptive routing control which improves the network performance. With the SA and SDN approaches applied by the framework, the total latency for transferring video streams can be minimized and the scalability of the CVM system can be significantly enhanced.

Keywords: cloud computing · video monitoring · video surveillance · software-defined networking

1 Introduction

A video monitoring or surveillance system has been applied for over a half-century [1] and employed for many applications such as home surveillance, traffic monitoring, transport safety, anomaly identification, and crime prevention. Video monitoring cameras deployed in the system have grown from the analog closed-circuit television (CCTV) to the Internet protocol (IP) cameras. In addition, a dashcam (or dashboard camera) which is used in the vehicle and smart glasses (e.g., Google Glass) which is a wearable display device with a built-in tiny camera and network connectivity can be potentially applied to the video monitoring system as well [2, 3].

Recently, cloud computing has been applied to a video monitoring system for providing value-added services [4]. In a cloud-based video monitoring (CVM) system, value-added services associated with on-demand storage capacity and scalable processing power are available for video distribution, video backup, video database management system, video encoding, and image/video recognition, for example. In practice, video data streamed from analog CCTV, webcams, and IP cameras can be transferred to the cloud services. Then, the video data processed or stored by the cloud services can be remotely monitored and managed by the users from their Internet-connected devices such as surveillance monitors, personal computers, smartphones, tablets, and set-top boxes. The CVM system can serve a large number of cameras and users. Dropcam [5], Smartvue [6], and Ivideon [7] are, for example, CVM providers.

In terms of client/server architecture, both cameras and users are *clients* of the cloud services. When numerous clients are simultaneously connected to the cloud services for sending and receiving video streams, scalability problems could incur although the CVM system provides a large pool of resources and network bandwidth to the server-side cloud services. Since the network for the clients is usually shared by multiple devices and users, a serious network bottleneck could easily incur in such an access network. When multiple clients operate in a congestive network with the small amount of Internet bandwidth (especially, bandwidth for home users or small enterprises), video streams from cameras or cloud services will cause a size scalability problem. The small amount of buffer memory allocated for a camera is another issue that increases jitter in streaming video. In addition, a large number of active clients located in different physical locations could increase the delay to transfer video streams, and then could cause a geographical scalability problem due to network delays.

To address the aforementioned scalability problems, this paper proposes a CVM framework. The framework applies two major approaches, namely *stream aggregation* (SA) and *software-defined networking* (SDN). The SA approach provides *aggregators* which are buffering servers with high speed network connectivity bridging clients and cloud services. In a specific physical area, an aggregator can be allocated close to a set of clients and can connect these clients together. Thus, the network delay can be significantly reduced by the aggregator. The scalability can be improved by the aggregator, since numerous traffic from the clients can be preprocessed, buffered, and cached to offload the network bandwidth and the cloud services. The aggregators can be offered to the CVM provider by Internet service providers (ISPs) or cloudlet providers [8].

The SDN approach offers programmability for network management in which network performance can be improved [9]. With the SDN approach, the CVM provider can flexibly manage a virtual network (consisting of virtual routers and links), and also allocate the appropriate amount of network bandwidth from ISPs to the aggregators and the cloud services. With the virtual network management, the CVM provider can design the optimal control of traffic scheduling, congestion, and routing. As a result, by applying both SA and SDN approaches together, the network latency can be minimized and the scalability of the CVM system can be significantly improved.

The rest of this paper is organized as follows. Related work is reviewed in Section 2. Section 3 describes the proposed CVM framework. Next, usage scenarios of the proposed CVM framework are demonstrated in Section 4. Finally, the paper is summarized in Section 5.

2 Related Work

Recently, video streaming services including video monitoring systems are trending to be deployed in a cloud infrastructure or platform since cloud computing offers a large pool of resources. Most resources can be provisioned on demand and charged on a pay-per-use basis. Cloud-based video monitoring systems have been designed and developed. For example, a video monitoring system proposed in [10] applied the Hadoop framework for storing and processing video streams in a large scale cloud computing infrastructure. In [11], the SmartHub system was proposed to process video streams from cameras in order to overcome limitations of the cameras. An algorithm was proposed in [12] to choose appropriate cloud providers for improving the quality of service (QoS) of video streaming and reducing the cost to utilize the cloud resources. In [13], the Scalable Video Coding proxy based on cloud computing was proposed that can deliver high-quality video streams. An integer programming model was formulated in [4] to allocate virtual machines (VMs) provided by a cloud provider for operating a video surveillance platform. The model minimizes the number of VMs, while QoS requirements of the platform can be fulfilled.

Although cloud computing provides a scalable computing environment for video streaming services, there are scalability problems when a large number of cameras in different locations simultaneously transfer video streams to cloud services via unreliable networks. The concept of aggregators or concentrators is introduced to deal with the problems. Aggregators have been applied in a smart (electrical) grid to achieve scalable connectivity that interconnects a large number of smart meters which simultaneously transfer data, related to metered power usage, to the same target [14]. In a video streaming system, the work in [15] applied multiple aggregators for the distribution of live multimedia content over wireless Internet connections in which the quality of video streaming can be improved.

Delivery of bulk video data from the growing number of cameras to cloud computing requires a large amount of network bandwidth and can incur a geographical scalability problem. To address such issues, the software-defined networking (SDN) approach can be applied to achieve better network performance. In [16], the OpenQoS controller design based on the SDN approach was proposed to enable the quality of service for multimedia delivery.

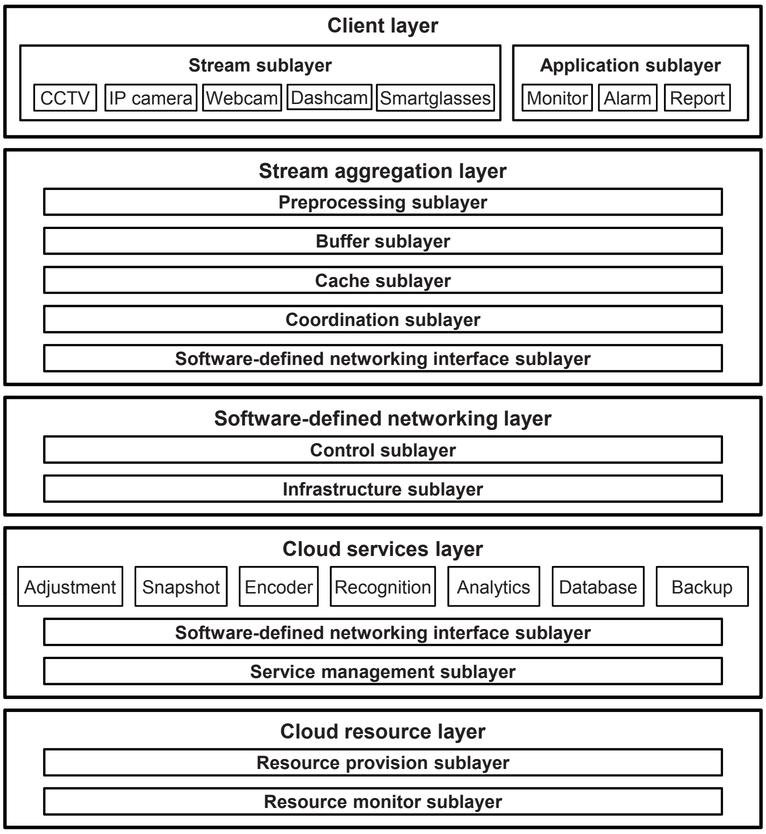

Fig. 1. Architectural layers of the proposed framework.

Unlike previous works, this paper proposes the cloud-based video monitoring framework that applies particularly both aggregators and SDN approaches to gain the benefits of the two approaches. That is, the aggregator approach can greatly lessen the network latency over the clients' network, while the SDN approach which provides the flexible network management can minimize the network latency in the network connectivity of both aggregators and public cloud providers. As a result, the size and geographical scalability of the cloud-based video monitoring system based on this framework can be significantly enhanced.

3 Proposed Framework

This section presents the proposed CVM framework primarily designed for a CVM provider. Fig. 1 depicts the architectural layers of the framework. The framework consists of five layers, i.e., client, stream aggregation (SA), software-defined networking (SDN), cloud services, and cloud resource layers. In particular, this paper introduces both SA and SDN layers that offer intermediate services for enhancing the size and geographical scalability of the CVM system. Without applying the SA and SDN services, the total delay in a traditional CVM system (denoted by d_{total}) for transferring a piece of video streams to or from the client is mathematically expressed as follows:

$$d_{total} = d_{client} + d_{service} + d_{network} \tag{1}$$

where client delay (d_{client}), service delay ($d_{service}$), and network delay ($d_{network}$) accumulate to render the total delay. For the camera, the client delay may include processing, buffering, and streaming delays in the camera. For the user who accesses an application of the CVM system, the client delay involves the time needed to load and execute the application and other related processes, and the time needed to render the video streams on the user's terminal screen, for example. The service delay includes the time needed to process functions of cloud services and access cloud resources. The client delay could be difficult to be reduced without upgrading the hardware of the client's terminal (e.g., cameras or the user's access device), while the service delay could be mitigated by software-engineering and high performance computing techniques. For example, programming code of a cloud service can be optimized by a compiler, tuned by a better software design, or parallelized by a programmer. In addition, a load balancing technique can be applied to efficiently distribute requests from clients to servers for gaining faster response time or greater throughputs of the services. The network delay includes several factors such as transmission and propagation delays, and queueing and processing delays in network equipment (e.g., routers and switches) [17].

The framework proposed in this paper mainly reduces the network delay by the services offered by the SA and SDN layers. The SA layer provides aggregators located in different physical locations. An aggregator acts as a consolidation hub of video streams close to cameras and users located in a specific boundary. The group of cameras located in the same proximity can be connected to the same aggregator via a high speed network connection that is much faster than the connection of the clients used to directly access to the cloud services. Moreover, the aggregator can include preprocessing functions for video streams before forwarding the streams to destined devices or cloud services for offloading the network bandwidth and cloud services. The aggregator can act as a cache server to store popularly accessed and recently received video streams which can additionally enhance the CVM system performance.

The SDN layer provides functions to dynamically adjust the network configuration according to the current network status and the amount of arrival traffic (e.g., data from video streams). The configuration includes routing, traffic scheduling, congestion control, and network bandwidth allocation. With an appropriate network configuration, the overall performance of the CVM system can be significantly improved.

The total delay in the CVM system based on the framework (denoted by d^*_{total}) for transferring a piece of video streams to or from the client is mathematically expressed as follows:

$$d^*_{total} = d_{client} + d_{service} + d_{aggre} + d_{sdn} \tag{2}$$

where d_{aggre} and d_{sdn} denote the delays incurred in the SA and SDN layers, respectively. With the optimal control of SA and SDN services, the delays in the SA and SDN layers can be greatly diminished (i.e., $d_{aggre} + d_{sdn} << d_{network}$). As a result, the total delay incurred in the framework can be optimized as well. The optimization approach to achieve the optimal control of SA and SDN services which is out of scope in this paper will be addressed and presented in the future work.

In Fig. 1, a layer provides a service to the adjacent layer above it and is served by its lower adjacent layer. Next, the five layers of the framework are discussed from the bottom layer to the top layer as follows:

3.1 Cloud Resource Layer

This cloud resource layer provides services to access a pool of cloud computing resources and services such as servers, networks, and storage for operating the cloud services located in its above layer. Basically, the CVM provider can rent resources from other public cloud providers such as Amazon, Microsoft, and Google to reduce the total cost of ownership (TCO). The CVM provider might operate local cloud resources in its private cloud. The CVM providers also can implement the hybrid cloud to gain a larger scalable computing infrastructure [18].

In Fig. 1, the cloud resource layer consists of two sublayers, i.e., resource provision and resource monitor sublayers. The resource provision sublayer is responsible to prepare and allocate cloud computing resources to the cloud services for accommodating requests and processing video streams received from the clients. For reducing the resource allocation cost, the resource provision sublayer should be able to resize the cloud computing resources to sufficiently meet the actual resource demand. Generally, most of the public cloud providers offer the on-demand provisioning option by which the CVM provider can increase or decrease the amount of cloud computing resources on demand. In some public cloud providers (e.g., Amazon and Microsoft), the reservation provisioning option is available which can significantly reduce the resource allocation cost for a long-term usage. Under uncertainties of resource demand from the clients and resource prices offered by the public cloud providers, the resource allocation cost can be efficiently minimized by the Optimal Cloud Resource Provisioning (OCRP) algorithm [19].

The resource monitor sublayer provides a service to monitor the cloud computing resources and get the last status of the resources. According to the monitoring status, the resource provisioning sublayer can dynamically adjust the amount of cloud computing resources. Cloud monitoring platforms and services that can be applied in this sublayer were discussed in [20].

3.2 Cloud Services Layer

The cloud services layer is the layer where different cloud services for the CVM system are available to the clients. These cloud services provide functions to process and store video streams. A number of video processing approaches which could be implemented in some cloud services were reviewed in [21]. Examples of the cloud services shown in Fig. 1 are listed as follows:

- *Adjustment service:* This service provides basic video processing functions such as noise reduction and contrast/sharpness/brightness adjustments.
- *Snapshot service:* A video stream can be transformed to a shorter video stream or a sequence of still images by this service given a specific range of the original stream. This service might be able to capture audio from the video stream as well.
- *Encoder service:* The encoder service provides video and audio codecs [21] for encoding, respectively, video and audio from a video stream.
- *Recognition service:* The recognition service provides functions to identify objects in a video stream. Objects appeared in video streams which could be detected by this service are, for example, human faces [22], animals [23], vehicle license plates [24], and events (e.g., fire and flood [25]).
- *Analytics service:* The (video) analytics service mainly provides surveillance functions to automatically detect events. This service could be applied for detecting and preventing accidents, crimes, and terrorism [26]. Video anomaly identification approaches to detect suspicious activities for addressing security threats which could be implemented in this service were discussed in [27]. The service could be also applied for marketing and other business intelligence [28].
- *Database service:* The video database management system which is specifically designed for organizing and retrieving video data [29] can be offered by this service.
- *Backup service:* This backup service provides backup and recovery functions to maintain video streams in cloud storage.

In Fig. 1, this cloud services layer consists of SDN interface and service management sublayers. The SDN interface sublayer is responsible to interact with services offered by the SDN layer. The service management sublayer provides services for distributing requests (i.e., video streams from the cameras or service requests from the applications) to and monitoring status of the cloud services. According to the status of the cloud services, the cloud computing resources in the cloud resource layer can be efficiently provisioned.

3.3 Software-defined Networking Layer

The software-defined networking (SDN) layer provides functions based on the SDN approach [9]. This SDN approach decouples control and data planes and provides programmability in the control plane. Hence, this approach allows application developers to simply and flexibly manage the control plane, which is related to the routing control of data packets, from their applications. With the SDN approach, the network configuration can be programmatically controlled and dynamically optimized, and the network performance can be improved [9].

Basically, the SDN approach consists of three layers, namely infrastructure, control, and application layers. The infrastructure layer which is the lowest layer consists of switching devices (e.g., routers and switches). These devices operate in the data plane, and provide functions to process and forward packets. They could also provide functions to collect and report network status. The control plane provides services (for example, as the form of application programming interface or API) for accessing functions of the switching devices. The functions can be provided for traffic scheduling, packet routing control, and congestion control. The application layer is the top layer where SDN applications access and control the switching devices by invoking the services in the control layer. In Fig. 1, the SDN layer of the proposed framework consists of the control and infrastructure sublayers equivalent to the two layers in the basic SDN approach, while the application layer is equivalent to the SDN interface sublayers existing in the cloud services and stream aggregation layers.

The control sublayer in this SDN layer can provide services for adaptive routing and virtual network management which are controllable by the SDN interface sublayers. Hence, the SDN interface sublayers can retrieve the last status of the virtual network such that appropriate traffic scheduling, congestion, and routing controls, which could significantly reduce the network latency and increase the system scalability, can be made for transferring video streams or service requests.

In the control sublayer, ISPs and network operators could offer reservation and on-demand options to the CVM provider for allocating network bandwidth to both cloud services and aggregators. Network bandwidth allocated by the reservation option is generally cheaper than that of the on-demand option for the long-term utilization. With the reservation option, the amount of bandwidth needs to be purchased in advance and usually cannot be refunded. In contrast, the amount of bandwidth can be dynamically provisioned at the moment when the bandwidth is needed with the on-demand option. While network bandwidth demand and price might be uncertain, the reserved bandwidth could be either overprovisioned or underprovisioned. Furthermore, the higher cost of the on-demand option will be unavoidable when the reserved bandwidth is underprovisioned. The optimization approach proposed in [30] to address the SDN-based network bandwidth allocation can be applied, while the two bandwidth allocation options and the uncertainties are taken into account.

3.4 Stream Aggregation Layer

The stream aggregation (SA) layer provides a number of aggregators with high speed network connectivity bridging clients and cloud services. Basically, an aggregator acts as a proxy server close to a group of clients located in the same proximity such that the network latency between clients and cloud services can be greatly diminished.

Aggregators can be provided by ISPs to the CVM provider. In addition, the concept of cloudlets can potentially form an infrastructure for aggregators. Originally, the term *cloudlet* coined in [8] is a resource-rich server or computer cluster that serves nearby mobile devices for mobile computing. The mobile devices could be directly connected to the cloudlet via either wireless telecommunications (e.g., 3G/4G) or wireless LAN (e.g., WiFi) connectivity. The cloudlet can synchronize data with and utilize larger resources from public cloud providers (e.g., Amazon and Microsoft) via high speed WAN connectivity. With this concept, a cloudlet provider could be available to operate computer clusters or micro-datacenters located in different physical locations. Resources from the cloudlet provider can generally serve other applications as well, not just mobile computing. The CVM provider can rent the resources including servers and storage from multiple cloudlet providers for operating the aggregators.

In Fig. 1, the SA layer contains five sublayers as follows:

- *Preprocessing sublayer:* In this sublayer, the aggregator may provide services for preprocessing functions such as video encoding and adjustment to offload both network bandwidth and cloud services.
- *Buffer sublayer:* This sublayer provides queue management system for buffering video streams that can avoid clogging networks interconnecting with cloud services and clients.
- *Cache sublayer:* The aggregator can store frequently accessed and recently received video streams with cache management system provided by this sublayer such that the network latency for accessing the video streams from the cloud computing resources can be significantly lessened. This cache management system might not be available for, especially, security monitoring systems because of some security matters. That is, video streams which are recently received or monitored by a user could be wiped from an aggregator since the streams might contain some sensitive information. In this paper, this cache sublayer is introduced for some other applications, for example, home entertainment and social networking applications. That is, online high-definition movies and video clips shared by social networking users can be managed by the cache management system for better video streaming.
- *Coordination sublayer:* The coordination sublayer provides a service in which the aggregator can interconnect with other aggregators for sharing necessary information among the aggregators. The shared information obtained from this interconnection service is used by the cache management system (in the cache sublayer) and can improve the performance of the CVM system in some situations, for example, the situation when video streams stored in one aggregators will be soon accessed by surveillance monitors connected to other nearby aggregators and the situation when a portable camera (e.g., dashcam or smartglasses) is transferring video streams to a sequence of aggregators by which the camera will pass.
- *SDN interface sublayer:* This SDN interface sublayer provides a service for the aggregator to interact with the SDN layer.

3.5 Client Layer

The client layer is the layer closest to the cameras and users. In Fig. 1, the client layer is divided into stream and application sublayers which are independent of each other. The stream sublayer is the layer where cameras capture and transfer video streams to aggregators. Cameras could be CCTV, IP cameras, webcams, dashcams, and smartglasses. The application sublayer is the layer where applications utilized by the end user are operated. Examples of the applications shown in Fig. 1 are listed as follows:

- *Monitor:* The monitor application can be used as monitoring and surveillance tools. With the value-added cloud services, the application could detect objects, human faces, or interesting events. The application could also distribute video streams encoded in various formats to fit the performance of the user's terminal device.
- *Alarm:* The alarm application can notify all devices that subscribe to specific events. For example, a surveillance camera detects a suspicious person in a restricted area, and then sends an email regarding this detection to all subscribed email addresses.
- *Report:* The report application could be used to produce a summary report about video streams, cameras, or system usage, e.g., the total size of video streams which were recorded and transferred, the current number of malfunctioning cameras, and the number of active users. The report might include the summary of processing results from the recognition and analytics services such as the list of physical areas where anomaly events were identified in the last month and the number of interesting objects found in a specific day and location.

3.6 Overview of Framework Implementation

An approach recommended by this paper to implementation of the proposed framework for a CVM provider is overviewed in Fig. 2. In this approach, a set of cloud computing resources including storage and VMs (as servers) can be provisioned from public cloud providers such as Amazon Web Services (AWS) and Microsoft Azure to operate the cloud services. The number of provisioned VMs should be dynamically resized according to the current load of the VMs and arrival of video streams. The public cloud providers can operate their resources in datacenters located in different geographical locations. The CVM provider should provision the resources from locations close to the clients. For example, Singapore, Ireland, and Oregon shown in Fig. 2 are locations where the CVM provider's clients are served by the CVM system.

ISPs or WAN operators can provide SDN-based virtual networks to the CVM provider. Traffic scheduling, routing and congestion controls can be flexibly managed to fulfill QoS requirements of the CVM system. Aggregators can be provisioned by cloudlet providers. In Fig. 2, a group of clients, i.e., users and cameras in houses, commercial buildings, and vehicles, located in the same proximity can

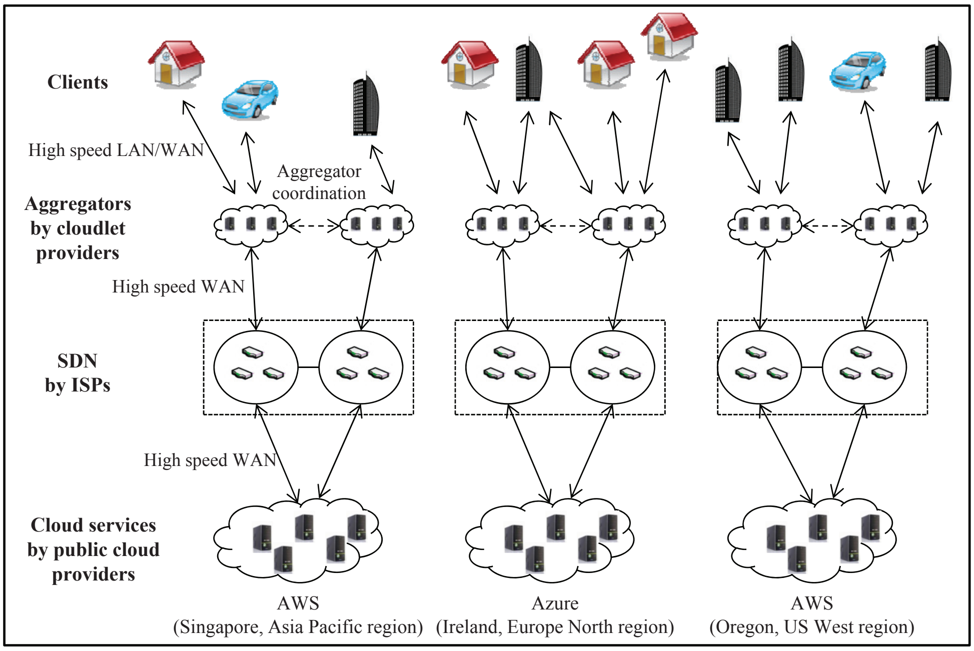

Fig. 2. Approach to framework implementation.

be connected to the same aggregator. Some clients could be connected to multiple aggregators (i.e., multihoming aggregators) to increase the reliability of the network connectivity. Fig. 2 shows that the aggregators require the service of the coordination sublayer to exchange information about movable clients (e.g., vehicles) so that the cache management system can be managed dynamically.

Regarding the network connectivity, the high speed LAN and WAN should be allocated to the link between clients and aggregators (e.g., gigabit Ethernet LAN or FTTx for the stationary clients and WiFi or LTE network for the movable clients). The high speed WAN is required for the connection between aggregators and public cloud providers.

4 Usage Scenarios

In this section, two different usage scenarios applying the proposed CVM framework are demonstrated as follows:

4.1 Home Security System

Currently, a CVM system has been applied in the home security system. Value-added services such as motion detection and face detection provide useful information to home users. Dropcam and Ivideon are, for example, CVM providers offering services for the home security system. However, both Dropcam and Ivideon might not be able to serve numerous users and handle video streams from numerous cameras in some locations where network latency in the network between the clients (i.e., the users and the cameras) and the providers' datacenters are too high. Due to such geographical and size scalability problems, home users might not be satisfied with the system. The proposed CVM framework could be applied in a home security system to improve the overall service performance and also provide a variety of value-added services offered by the cloud services such as fire and flood detection and anomaly identification at home.

4.2 Traffic Monitoring System

A traffic monitoring system can gain benefits from the proposed CVM framework. In this system, a set of CCTV cameras are located in different locations for monitoring streets. The system could be used to trace events on the streets (and their nearby places) and control the traffic. Video streams from the CCTV cameras are consolidated by the aggregators before being sent to cloud services by the SDN. The traffic police and staffs from the traffic authority can get useful information from the cloud services such as accident detection, face detection, license plate detection, and traffic jam prediction for better safety and traffic management.

The traffic authority might invest and deploy the aggregators in this system. The aggregators should be located in different physical areas to cover a group of CCTV cameras which are in close proximity. Then, the SDN service offered by ISPs provides better network performance for delivering the video streams among traffic control centers of the authority and the cloud services.

5 Conclusion

Cloud computing has been applied to a video monitoring system to offer a large pool of storage and processing power for storing and processing video streams captured by cameras. However, the size and geographical scalability problems could be encountered when a large amount of video streams is transferred to or from the cloud. To address the problems, the CVM framework has been proposed in this paper. Two main approaches have been applied by the framework, i.e., stream aggregation (SA) and software-defined networking (SDN). The SA approach can significantly reduce the network latency by providing a number of aggregators close to the clients (i.e., users and cameras). Furthermore, the cache management system and preprocessing functions of the aggregators can even greatly diminish the overall network latency in the clients' network connectivity. The SDN approach for the flexible network management can minimize the network latency in both the clients' and the aggregators' network connectivity. By applying both SA and SDN approaches together, the size and geographical scalability of the CVM system can be significantly enhanced.

For the future work, optimization approaches to achieve the optimal control of SA and SDN services (including cloud computing resources and network bandwidth provisioning, cache management, traffic scheduling, routing and congestion controls) will be studied. A prototype of the proposed framework will be implemented and evaluated with the use of discrete event simulation. Then, we will evaluate and analyze performance matrics (e.g.,, network latency and service delay) of CVM systems with and without the framework.

References

1. Williams, C. A.: Police surveillance and the emergence of CCTV in the 1960s. Crime Prevention & Community Safety, 5(3), 27–37 (2003)
2. Hildebrandt, P.: Dash-Cams Keep Record: Recording Officers' Interactions with the Public with Mobile Video Isn't Enough, SOPs Must Clarify How Video is Captured and Stored. Law Enforcement Technology, 36(2), 10–14 (2009)
3. Kurze, M., Roselius, A.: Smart glasses linking real live and social network's contacts by face recognition. In Proceedings of the 2nd augmented human international conference (p. 31). ACM (2011)
4. Hossain, M. S., Hassan, M. M., Qurishi, M. A., Alghamdi, A.: Resource allocation for service composition in cloud-based video surveillance platform. IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 408–412 (2012)
5. Dropcam, `http://www.dropcam.com/`
6. SmartVue, `http://www.smartvue.com`
7. Ivideon, `http://www.ivideon.com/`
8. Satyanarayanan, M., Bahl, P., Caceres, R., Davies, N.: The case for vm-based cloudlets in mobile computing. Pervasive Computing, IEEE, 8(4), 14–23 (2009)
9. Xia, W., Wen, Y., Foh, C. H., Niyato, D., Xie, H.: A Survey on Software-Defined Networking. Communications Surveys & Tutorials, IEEE (2014)
10. Lin, C. F., Yuan, S. M., Leu, M. C., Tsai, C. T.: A framework for scalable cloud video recorder system in surveillance environment. In Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on, 655–660 (2012)
11. Saini, M. K., Atrey, P. K., Saddik.: From Smart Camera to SmartHub: Embracing Cloud for Video Surveillance. International Journal of Distributed Sensor Networks (2014)
12. Chen, W., Cao, J., Wan, Y.: QoS-aware virtual machine scheduling for video streaming services in multi-cloud. Tsinghua Science and Technology, 18(3), 308–317 (2013)
13. Huang, Z., Mei, C., Li, L., Woo, T.: CloudStream: Delivering high-quality streaming videos through a cloud-based SVC proxy. In INFOCOM, 2011 Proceedings IEEE, 201–205 (2011)
14. Wu, G., Talwar, S., Johnsson, K., Himayat, N., Johnson, K. D.: M2M: From mobile to embedded internet. Communications Magazine, IEEE, 49(4), 36–43 (2011)
15. Kamilova, M. I., Hesselman, C., Widya, I., Huizer, E.: Adding policy-based control to mobile hosts switching between streaming proxies. In Policies for Distributed Systems and Networks, 2005. Sixth IEEE International Workshop on, 243–246, (2005).
16. Egilmez, H. E., Dane, S. T., Bagci, K. T., Tekalp, A. M., Koc Univ., Istanbul, Turkey. In Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific, 1–8. IEEE, (2012)

17. Kurose, J. F., Ross, K. W.: Computer networking: a top-down approach. 6th edition, Pearson Education, (2013)
18. Sotomayor, B., Montero, R. S., Llorente, I. M., Foster, I.: Virtual infrastructure management in private and hybrid clouds. Internet Computing, IEEE, 13(5), 14–22 (2009)
19. Chaisiri, S., Lee, B. S., Niyato, D.: Optimization of resource provisioning cost in cloud computing. Services Computing, IEEE Transactions on, 5(2), 164–177 (2012)
20. Aceto, G., Botta, A., De Donato, W., Pescapè, A.: Cloud monitoring: A survey. Computer Networks, 57(9), 2093–2115 (2013)
21. Dasu, A., Panchanathan, S.: A survey of media processing approaches. Circuits and Systems for Video Technology, IEEE Transactions on, 12(8), 633–645 (2002)
22. Connolly, J. F., Granger, E., Sabourin, R.: An adaptive classification system for video-based face recognition. Information Sciences, 192, 50–70 (2012)
23. Burghardt, T., Ćalić, J.: Analysing animal behaviour in wildlife videos using face detection and tracking. IEE Proceedings-Vision, Image and Signal Processing, 153(3), 305–312 (2006)
24. Du, S., Ibrahim, M., Shehata, M., Badawy, W.: Automatic license plate recognition (ALPR): a state-of-the-art review. Circuits and Systems for Video Technology, IEEE Transactions on, 23(2), 311–325 (2013)
25. Lai, C. L., Yang, J. C., Chen, Y. H.: A real time video processing based surveillance system for early fire and flood detection. In Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE, 1–6 (2007)
26. Regazzoni, C. S., Cavallaro, A., Wu, Y., Konrad, J., Hampapur, A.: Video analytics for surveillance: Theory and practice [from the guest editors]. Signal Processing Magazine, IEEE, 27(5), 16–17 (2010)
27. Saligrama, V., Konrad, J., Jodoin, P. M.: Video anomaly identification. Signal Processing Magazine, IEEE, 27(5), 18–33 (2010)
28. Shan, C., Porikli, F., Xiang, T., Gong, S. (Eds.): Video Analytics for Business Intelligence (Vol. 409). Springer (2012)
29. Ardizzone, E., La Cascia, M.: Automatic video database indexing and retrieval. Multimedia tools and applications, 4(1), 29–56 (1997)
30. Chase, J., Kaewpuang, R., Wen, Y., Niyato, D.: Joint virtual machine and bandwidth allocation in software defined network (SDN) and cloud computing environments. In Proceedings of IEEE ICC, Sydney, Australia (2014)