International Journal of Combinatorial Optimization Problems and Informatics, 17(2), 2026, 357-377. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2026.v17i2.1168

www.editada.org

Content Web Services Deliverability: A Systematic Review

Sócrates Benítez-Domínguez, Olivia Graciela Fragoso-Diaz, Juan Carlos Rojas-Pérez
1 National Center for Research and Technological Development (TecNM/Cenidet)
E-mails: d16ce073@cenidet.tecnm.mx, olivia.fd@cenidet.tecnm.mx and juan.rp@cenidet.tecnm.mx

Abstract. Content Web Services are Web services whose main function is to deliver files or content. However, deliverability is difficult to measure if methods do not exist or if the attributes involved, such as those related to clients and servers, are not identified. To find the attributes involved, a systematic review was carried out with studies from 2015 to 2024 using the PRISMA method. In the first step, there were 2,660 studies from four digital libraries, after five inclusion and eight exclusion criteria, 79 studies were selected. Most of the works were found to be dealing with improving service architectures or computer networking. The 79 studies were fully read, four studies were finally selected, and only one deals with deliverability. There are platforms and attributes that influence content delivery; however, studies considering Web services for content delivery are very scarce.

Keywords: Systematic review; Web service; content; deliverability; capabilities; server attributes; client attributes

Article Info
Received July 3, 2025
Accepted Dec 11, 2026

1 Introduction

This study aims to identify relevant methods and their attributes used for content delivery through Web services (WS), which will be called content Web services (CWS). This study allows identifying the attributes of the servers and clients considered by different research studies that are involved in content delivery via WS.

Content may be represented by a file, such as PDF, Power Point, Word, Excel, image, audio, or video, which is delivered in response to a request. The content of the WS is encoded in BASE64, and according to (Grigorik, 2013) this encoding increases the size more than 30% and impacts computational resources, more disk space, memory space and CPU processing are required, so it also affects network bandwidth or network infrastructure. The volume of content may be greater than a transaction, a database operation, or another type of processing; therefore, the processing load may be greater for a CWS than for a transactional WS. Transactional WS deliver information about the processing of information such as database queries, algorithm execution, or another type of processing. For the reasons mentioned above, it is important to know what are the attributes that should be considered when delivering content, in addition, the success of content platforms relies on the performance of their architectures, so it is important to identify the attributes that influence the content delivery capacity through Web Services. It is necessary to know which attributes impact the proper functioning of Web Services, which attributes are adverse.

Multiple domains benefit from defining the attributes that must be considered when delivering content through Web Services. In addition, delivery capability has not been researched, evaluated, or given the importance it deserves. Even more, there is a lack of studies that evaluate the delivery capacity considering the service architectures. Content-based Web services are used in domains such as E-Learning, healthcare, cartography, geology, astronomy, multimedia, digital distribution platforms, web pages, websites, or applications where static content is separated from dynamic content, or where content such as documents need to be managed. In this context, the benefits of using static content and delivering it through Web services in the software engineering field are numerous. The main benefits of this approach include simplified application maintainability; the decoupling of components facilitates changes, updates, or corrections; modularity reduces application complexity and increases the reusability of its components; application efficiency is also improved by reducing processing and the use of computational resources; in addition, it shortens response times when content is delivered via Web services. Furthermore, it allows CWSs to be deployed independently, with each one being responsible solely for a single piece of content.

Delivery methods such as streaming, broadcasting, WebRTC, IPTV, etc., and methods focused on computing networks for content delivery have also been studied. The best known of the methods focused on computing networks are Network Virtualization Functions (VNF), Software-Defined Networking (SDN), Content Delivery Networks (CDN), etc. These are beyond the scope of this study because the purpose is to identify the attributes involved in the content delivery process through WS. It is important to emphasize that this study does not deal with content-related aspects, but instead concentrates on the underlying delivery mechanisms and infrastructures.

The remainder of this paper is structured as follows. Section 2 presents the theoretical foundation, including the description of SOA, Web services and an analysis between SOAP and REST. Section 3 details a resume of related SLR's discovered in the SLR process. Section 4 presents the SLR protocol and methods; the results are discussed in Section 5. Section 6 presents the research directions and current trends in service computing. Section 7 presents threats to validity, and finally, Section 8 produce the conclusions.

2 Background

Service-Oriented Architecture (SOA) uses WS as a method of communication between different applications and computers through a network. A Web Service is a program with a well-defined interface that can be published, located and invoked by using standard Web protocols (De Renzis et al., 2017). To date, there are two methods for WS implementation, SOAP and REST. SOAP is a protocol that provides information to the user and describes the operations and messages of the service (Kumari & Rath, 2015). REST stands for "Representational State Transfer" and is an abstraction of the architectural elements within a distributed hypermedia system and is composed of a set of principles that define the interaction between different components. SOAP and REST have been compared in several studies. According to (Tihomirovs & Grabis, 2016) (Pautasso et al., 2008), SOAP-based WS should be implemented for business (B2B) projects. Table 1 provides a comparative summary of the advantages and disadvantages of SOAP-based and REST-based web services, according to authors (Curbera et al., 2002; Soni & Ranga, 2019; Tihomirovs & Grabis, 2016).

Table 1. Advantages and disadvantages between SOAP and REST *(Curbera et al., 2002; Soni & Ranga, 2019; Tihomirovs & Grabis, 2016).*

WS Type	Advantages	Disadvantages
SOAP	Communication oriented model Higher security and reliability Lower number of errors Asynchronous requests Distributed computing Support from other standards and extensible (WSDL,WS,WS-Security, WS-Atomic Transaction)	More expensive and, consequently, less flexible and has lower performance due to its XML foundation. More LOC + XML Marshaling and unmarshaling processes
REST	Resources-oriented model Uses another data formats (JSON, HTML, XML) Greater scalability Compatibility Performance Simplicity	Use the existing standards Limited bandwidth

Table 2 describes the architectural differences between SOAP and REST according several authors, the main characteristics presented in this table are the structure of the message, the interaction and the architectural structure and the architectural style of REST composed of restrictions unlike SOAP which is composed of a set of rules that define how two entities communicate, known as protocol.

Table 2. Comparative summary of the architectural differences between SOAP-based WS and REST *(Bellido et al., 2013; Castillo et al., 2011; Curbera et al., 2002; Halili & Ramadani, 2018; Kumari & Rath, 2015; Pautasso et al., 2008; Soni & Ranga, 2019; Tihomirovs & Grabis, 2016).*

Attribute	SOAP	REST
Description	Protocol Simple Object Access Protocol, developed by IBM and Microsoft	Representation Style Transfer is an abstraction of the architectural elements within a distributed hypermedia system. Constraints: *Null Style* *Client-Server* *Stateless* *Cache* *Uniform Interface* *Layered System* *Conde-On-Demand*
Protocols	HTTP,SMTP, FTP	HTTP
Description Language	WSDL	WADL
Responses	XML	XML, JSON, HTML
Message Structure	SOAP Message: *Envelopment* *Header Body*	REST Message: *Transport* *Header* *Resource*
Interaction	RPC and SOAP Actions	HTTP Verbs
Structure	Transmitter Receiver Intermediary	Origin Server Gateway Proxy User Agent Connectors: *Client* *Server Cache* *Resolvers Tunnel*
Directory	UDDI	N/A
Network Layer	Transport	Application

One of the identified advantages of SOAP over REST is that it allows the implementation of two methods for sending files. The first is Web Services-Interoperability Attachments Profile (WS-I Attachments), which allows files to be attached to the web service response without being sent in the SOAP message. The second mechanism is Message Transmission Optimization Mechanism (MTOM), a newer and optimized version of WS-I. MTOM is defined as a method for transmitting attachments by converting Base64 Binary encoded data into raw binary data, thus improving performance and reducing transmission overhead.

On the other hand, microservices can be considered an evolution of service-oriented architecture (SOA). They are characterized as loosely coupled services, each oriented toward a specific business objective, and communicating with each other through well-defined interfaces. Furthermore, each microservice is developed, implemented, and maintained independently, allowing for greater scalability, resilience, and architectural flexibility.

Microservices have seen widespread adoption in areas such as the Internet of Things (IoT), where heterogeneous devices communicate using sensors and actuators. In this context, (Gaur et al. 2017) define IoT as “a network connecting uniquely identifiable objects to the Internet, which possess sensing/actuation capabilities and programmable potential.” This definition emphasizes the need for distributed and scalable architectures, where microservices have become established as a suitable option for managing and orchestrating IoT devices.
With the expansion of the Internet, a growing interest emerged in the exchange of static resources, such as images and files, distributed through web protocols and peer-to-peer architectures. This interest enabled the development of various content delivery strategies. For example, Content Delivery Networks (CDNs), one of the pioneering projects in this field being Akamai,

launched in 1999 by Thomson Leighton (Dilley et al., 2002), marked the beginning of the CDN revolution. To date, it is one of the most widely used strategies, to which components such as container technologies; software-directed networking functionalities; communication protocols like HTTP, HLS, and WebRTC; and various types of algorithms for packet routing, cache management, and content optimization have been added, among others, with the aim of improving the quality of content delivery.

Streaming is one of the most widely used methods for transmitting content over the internet. Developed by RealMedia (Heller, 1998) as a strategy for transmitting audio and video, YouTube capitalized on this method in 2005 as a content-sharing platform (Cheng et al., 2007). Two years after its launch, it already accounted for 10 percent of internet traffic. Meanwhile, the MPEG-DASH (Dynamic Adaptive Streaming over HTTP) standard, introduced in 2011, enables the adaptive streaming of multimedia content without requiring codecs or formats for proper display in web browsers. This advancement led to other improved solutions, such as HTTP Live Streaming (HLS).

Since then, content delivery strategies have continued to evolve, allowing users to access higher-quality multimedia content with lower latency and an improved user experience.

3 Related Reviews Discovered in the SLR Process

In the analysis of this systematic literature review (SLR), no SLR's, surveys or systematic mappings related for content delivery via WS were found. However, some reviews about topics of interest can be considered. Resource allocation is a recurring topic, in (Jyoti et al., 2020) the authors perform a comparative study of load balancing algorithms to determine, illustrate, compare, and analyze the newest methods developed for load balancing and service brokering, considering dynamic resource allocation and cost reduction in Cloud Computing. In the same area of resource allocation, (Hani et al., 2015) and (Faniyi & Bahsoon, 2016) analyze Service Level Agreement (SLA) to increase flexibility and scalability. In (Mustafa et al., 2015) the authors perform a comprehensive review of resource management techniques, in which a taxonomy based on different characteristics is developed.

Another recurring topic of interest is improving network performance; in (Barakabitze et al., 2020) the authors present a review and updated solutions related to the 5G network slice, using software defined networking (SDN) and virtualization of network functions (NFV). In (Al-Turjman, 2020) the author produces a survey investigating heterogeneous networks of ultra-wide range, associated with a Big Data project and 5G-based networked IoT. In (Khan et al., 2021) the authors provide an analysis of the gaps in academia and industry of existing SOA-based SDN quality of service effectiveness management techniques to ensure service delivery. In (Xavier & Kantarci, 2018) the authors argue about studies on architectural design issues, virtualization solutions, and challenges in cloud communications and networking. In (Bhamare et al., 2016) the authors discuss open research topics related to software function chain (SFC) architecture and demonstrates the need for an analytical model for such an architecture in order to obtain optimal performance in Cloud Computing. In (Abdelmaboud et al., 2015) the authors review quality of service approaches in Cloud Computing to identify where more emphasis should be placed in current and future research directions. In (Le Duc et al., 2019) and (Su et al., 2016) the authors discuss studies on quality of experience (QoE) in video streaming, whereas (Aksakalli et al., 2021) the authors focus on microservices deployment approaches and communication platforms. Lastly, in (Asghari et al., 2019) the authors categorize research techniques on IoT applications. None of the previous studies analyze content delivery using Web services.

Within the health domain, several studies also address content delivery. For example, the work of (Roberts et al., 2018) analyzes the use of multimedia content for educational, instructional, and support purposes for individuals undergoing pulmonary rehabilitation. Similarly, (Kamali et al., 2020) present a SLR on virtual coaching systems focused on well-being and health within the eHealth field, integrating multiple types of content such as physiological monitoring, voice messages, and video calls. Likewise, (Baxmann et al, 2025) offer an SLR study on the management and adoption of new orthodontic techniques, highlighting the role of digital content in improving both clinical and educational processes.

In another relevant domain, the studies by (Li et al., 2024) and (Enríquez et al., 2022) describes the use of interactive exhibits didactic demonstrations, augmented and virtual reality, guided tours, and art presentations as mechanisms for enriching the learning experience. Additionally, the studies by (Shen & Li 2021) and (Chand & Ogul 2020) present SLR´s centered on content delivery.

4 SLR Protocol

This SLR is based on (Kitchenham & Charters, 2007) and it is used as a guide to describe Section 4.1 to Section 4.11

4.1 Research Question

This SLR has a single research question (RQ):

"What are the attributes involved in the process of delivery content through WS?"

4.2 Search Strategy

The digital libraries listed in Table 3 are considered in the search strategy. SLRs are not considered because they are secondary studies, experimental SLRs are excluded because they show evidence only from primary studies.

4.3 Search String

The search string was organized with keywords using the OR operators for synonyms or alternatives, AND to combine the keywords and NOT for exclusions or negations.

("qos" or "quality of service") and ("deliverability" or "delivery" or "Provisioning") and ("web service") and ("resources" or "content")

4.4 Related Work Sources

Table 3 shows the databases, digital libraries, or related work sources used in this SLR. The table also shows the sources ~~access~~ URL, the number of studies found in each of them, and the total number of studies obtained from the overall search.

Table 3. Digital libraries of scientific publications.

Source	URL	Studies
ACM Digital Library	https://dl.acm.org/	388
ScienceDirect	https://www.sciencedirect.com/	796
IEEE Xplore	https://ieeexplore.ieee.org	1197
SpringerLink	https://link.springer.com	279
Total:		**2660**

4.5 Keywords:

The keywords *Web service, quality of service, delivery, content, and resources* were identified for this SLR, and Table 4 presents these keywords along with their synonyms and corresponding word combinations.

Table 4. Keywords, synonyms and word combinations.

Keyword	Synonym
Content	Resource
Delivery	Deliverability
Delivery	Provisioning
Quality of Service	QoS

4.6 Criteria Selection

For the purpose of this work, five inclusion criteria and eight exclusion criteria for the SLR were defined and listed below.

Inclusion criteria:

- Conference papers.
- Studies in the form of a scientific paper.

- Studies focused on delivering content through WS.
- Studies published between 2015 and 2024.
- Readable studies.

Exclusion criteria:

- Studies focused on computer networks.
- The full-text does not give sufficient details.
- Studies that have an extended version and produce the same intervention with more details.
- The document is a survey, systematic review, or similar.
- The full text of the document could not be obtained.
- Studies that do not have a clear description of their process or their interpretation is difficult.
- Studies that implement SOA architectures for activities other than content delivery through WS.
- Written in a language other than English.

4.7 Selection Procedure

PICO (Kitchenham & Charters, 2007) criteria were used to identify all studies dealing with the delivery of content through WS. To formulate the research question, the PICO technique was reviewed and applied as follows:

- Population (Problem Definition): Published studies that deal with delivering content through WS, other than traditional WS or transactional WS.
- Intervention (by keywords): attributes that intervene in the process of content delivery through WS.
- Control: Inclusion and exclusion criteria.
- Outcome: List of elements involved in content delivery through WS.

4.8 Quality Assessment

Given that databases and studies are recognized within the scientific community, the studies considered for this SLR are therefore considered quality works. However, an evaluation of selected studies was carried out under the following guidelines: context, readability, experimentation, and analysis of the experiment. The results validation of the selected studies was performed considering procedures, test scenarios, variables and metrics, and statistical methods. The origin of the test data was evaluated; three scales were defined in the quality evaluation:

- For simulated data, the quality is low.
- For mixed data, the quality is medium.
- For real data, then the quality is high.

4.9 Data Extraction

The data was collected by a single person using the MS Excel tool. The information is registered in different columns and spreadsheets. Data was extracted by answering a set of questions of each study.
The questions discussed in this SLR are as follows:

- Is the study about content delivery?
- Does the study use WS for content delivery?
- What is the application area of the study?
- What is the purpose of the study?
- What are the attributes considered for content or resources delivery?

Figure 1 illustrates the SRL selection process. Data extraction was carried out in two phases following the PRISMA guidelines. In step 1, questions 1 and 2 were answered by analyzing the abstract. From the 2660 studies, 69 were identified as duplicates. In Step 2, questions 3, 4 and 5 were answered using complete studies, which resulted in considering 79 of them. Finally, only four studies were identified as relevant and selected for reporting.

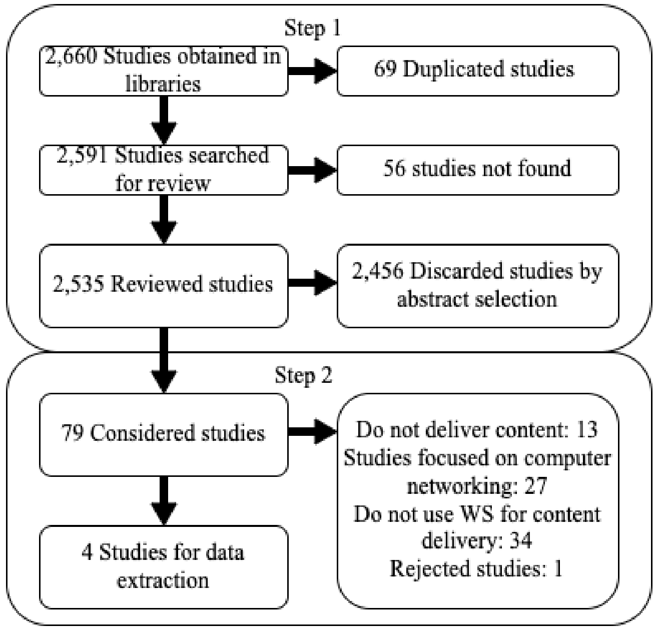

Figure 1. Systematic literature review selection process.

4.10 Data Synthesis

The approach is to find the attributes for content delivery and which are the most appropriate parameters for their use. Studies that focus on computer networks are discarded because their purpose is beyond our scope.

4.11 Execution

The execution process was carried out in four steps due to some difficulties in reaching data sources during the COVID19 pandemic. The first step was carried out in November 2021, with 678 studies found. In February 2022, the second step began with 645 studies found. The third step was carried out in January 2025, and to date the result was 1,337 studies found. At the end of the three steps, 2,660 studies were collected.

5 Results

This section summarizes the data extracted from the selected studies. Figure 2 displays the search results by year of publication, in this case, in the years 2015 to 2017 the volume of studies is greater than in the years 2018 to 2024, due to the attention paid to works on Cloud computing.

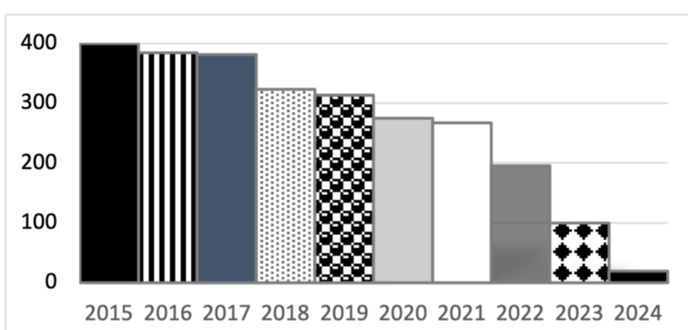

Figure 2. Studies found by year.

5.1 Approaches

From the studies found, three study areas were identified, as illustrated in Figure 3. Cloud computing is the most popular topic, followed by computer networking and finally, IoT.

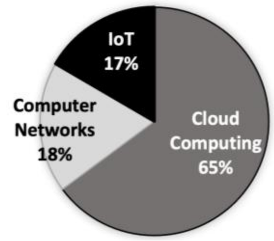

Figure 3. Retrieved studies by approach.

5.2 Delivery Methods Answer to Research Question

In step 2, in accordance with the exclusion criteria, 77 studies were excluded from the studies chosen for a complete full-text reading. Among the relevant topics of exclusion are: studies focused on computer networks, studies that do not deliver content, and use methods other than WS for content delivery. Figure 4 shows the percentages of the methods used for content delivery in the studies considered, Streaming, WebRTC and IPTV. The total of these are 45 studies, of which 5% are focused on IPTV, 11% are focused on Real Time Communication (RTC) and the rest implement or analyze the streaming or broadcasting method. Streaming is an event-based technology used to send data over a long-lived open communication channel. While Web services are software components designed to establish communication between two applications, and the goal is to establish interoperability, reusability and integration. There are three approaches for Streaming: Live Streaming, peer to peer, and broadcasting.

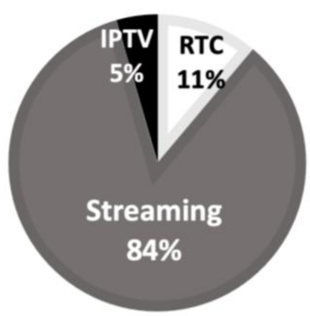

Figure 4. Percentage of methods used for content delivery in considered studies.

5.3 Excluded Studies on Computing Networks

Table 5 lists computing networks studies excluded and therefore considered outside the scope or intended purpose of this SLR as mentioned above. Column one provides the reference of each study, column two indicates the network method or infrastructure used to improve content delivery (e.g., VNF, SDN, CDN, DDS, etc.), and the third column specifies the type of content delivered in each referenced study. These studies also show that much of the research related to content delivery concentrates on improving data transmission across the network infrastructure, whether the networks are mobile, wired, or wireless, and regardless of the proximity of data centers to end users. On the other hand, Content Delivery Networks (CDNs) represent computational architectures designed to distribute content through mechanisms such as caching servers, load balancing, and fault tolerance; however, their emphasis remains on optimizing data transmission. These approaches typically rely on routing algorithms, traffic management strategies, flow scheduling, and other techniques.

Table 5. Excluded studies on computing networks.

Study	Network	Delivers
(Ahammad et al., 2021)	SDN, VNF	Video
(Aygun et al., 2021)	CDN	Video
(Ruan & Ye, 2019)	CDN	Video
(Wen et al., 2017)	CDN	Video
(Huang et al., 2016)	CDN, CCN	Multimedia
(Tian et al., 2016)	VNF	Video
(Zaher et al., 2021)	VNF	Data
(Z. Zhang et al., 2016)	VNF	Video
(Zhao & Medhi, 2017)	VNF	Streaming
(McCarthy et al., 2021)	ICN	Video
(Almadani, 2015)	DDS	Data Types
(Akhtar et al., 2015)	CDN, Hybrid-OSN	Multimedia
(Silva et al., 2015)	Heterogenous Networks	Information
(Abdou et al., 2015)	Mobile Networks	Messages
(Montella et al., 2017)	Mobile Networks	Data sets
(Chaudhuri, 2017)	Computer Networks	Data sets
(Xhagjika et al., 2017)	Cloud Distribution	WebRTC
(Sakthidasan et al., 2021)	WSN	Data
(Bhardwaj & Sharma, 2018)	WSN and Cloudlets	Data
(Quadri et al., 2018)	VNF	Video
(Wang et al., 2017)	CDN	Video streams
(De Rivera et al., 2016)	Pub/Sub Networks	Multimedia
(Suri et al., 2019)	Communication Protocol	Files
(Z. Wu et al., 2019)	CDN	Multimedia
(Ge et al., 2016)	WSN	Video
(Kilanioti & Papadopoulos, 2017)	CDN	Multimedia
(Zou et al., 2015)	WSN	Multimedia

5.4 Excluded Studies on Other Delivery Type

Table 6 presents the studies that use Web Services to manage applications but not for content delivery. These works propose methods aimed at improving delivery quality or strengthening the architecture or computational infrastructure. Among these studies, Almadani et al. (2015) introduces a prototype for transmitting measurement data during drilling operations, enabling real-time monitoring of downhole conditions through a flexible and fault-tolerant communication infrastructure.

Table 6. Excluded studies other delivery type.

Study	Purpose	Context
(Palumbo et al., 2021)	Resource Allocation	Cloud Service Provider
(de Assunção et al., 2016)	Elasticity	Cloud Computing
(Ketankumar et al., 2015)	Resource Allocation	Cloud Computing
(Fan et al., 2015)	User Personalization Framework	SaaS Service
(Gutierrez-Garcia & Sim, 2015)	Bag of Tasks Applications	Cloud Services
(Almadani et al., 2015)	Optimization of DDS	Computer Networks
(Jaiganesh et al., 2015)	QoS Evaluation	Cloud Services
(Gundu et al., 2021)	Load Balancing	Data Centers
(Gao et al., 2021)	Search Efficiency	IoT Environment
(Tellez et al., 2019)	Containerization	Edge Computing
(Savolainen & Rinne, 2015)	Collect Data	Service Catalogs
(Y. Wu et al., 2015)	Dynamic Scaling	Cloud Computing
(Gkamas et al., 2017)	Integration	Multimedia Apps

5.5 Excluded Studies on Content Delivery

Table 7 presents the list of studies that use web services to manage applications but do not employ them for content delivery. The proposed methods in these studies are focus on improving delivery quality or enhancing the architecture or computational infrastructure. It is important to note that many studies analyze data transmission metrics to determine appropriate strategies and to improve the infrastructure supporting content-based services, such as recommendation systems.

Table 7. Excluded studies on content delivery.

Study	Purpose	Delivery Method	WS Type
(Bhattacharyya et al., 2018)	Manage content transfer	CoAP Protocol	REST
(Bolettieri et al., 2021)	Resource allocation	Streaming	
(Shi et al., 2020)	Delivery evaluation model	Streaming	
(Zamani et al., 2020)	Resource allocation	Streaming	
(Santos et al., 2020)	Content orchestration	Streaming	
(Cánovas et al., 2018)	Intelligent media distribution	Video Streaming	
(Moon et al., 2016)	Video streaming model	Streaming	
(Rizvi et al., 2017)	SLA management model		
(Barba-Jimenez et al., 2016)	Model	Cloud VoD	
(Tudoran et al., 2016)	Middleware	Streaming	
(Bakri et al., 2015)	Communication protocol analysis	VR 3D	
(Hossain et al., 2017)	Collects data	Streaming	
(Wamser et al., 2016)	Algorithms	Service Selection	
(Holowczak & Houmansadr, 2015)	Unblock CDN		
(Oh et al., 2015)	Collaborative Platform	WebRTC	
(Sheltami et al., 2018)	Fog Computing	CoAP Protocol	
(Akpinar et al., 2017)	Fault Tolerance	WebRTC	REST
(Singh et al., 2020)	Agriculture Platform SaaS	Streaming	
(Farhad et al., 2017)	Video Transcoder	Streaming	
(Al Ridhawi et al., 2017)	Media Service Composition	Streaming	
(Alam et al., 2016)	Deliver Content	Streaming	
(Huf et al., 2016)	Deliver Content	Broadcast	SOAP
(Dantas et al., 2016)	Analytic Model	VoD Streaming	
(Seraoui et al., 2017)	Resource Allocation	IMS	
(Said et al., 2017)	Adaptation of Communication Protocol	RTC	
(Yamada et al., 2015)	RTC Platform	WebRTC	
(Tran et al., 2018)	Monitor Transmission	Streaming	
(Abu-Lebdeh et al., 2016)	Delivery Live Video	Streaming	
(Yoon, 2015)	Delivery Video	IPTV	
(Gaur et al., 2017)	Connect UAV's with IoT	Streaming	REST
(Sarker et al., 2019)	Multimedia Application	Streaming	
(Soltanian et al., 2018)	Web Service Composition	WebRTC	REST
(Cheng & Hancke, 2015)	Video Surveillance Management	Streaming	SOAP
(Rasbi & Singh, 2017)	Video	WebRTC	

5.6 Content Delivery Methods

Table 8 summarizes the differences among the different methods for content delivery examined in this study, based on contributions from several authors (Apple Inc., 2023; Cánovas et al., 2018; Moon et al., 2016; Seraoui et al., 2017; Stockhammer, 2011; W3C, 2023; Yoon, 2015). This table compares each method in terms of architecture type (centralized, decentralized, or peer-to-peer), communication model, transmission approach (multicast or unicast), distribution mode (on-demand or live), and the communication protocols employed for content delivery.

Content WS can be invoked from anywhere and at any time (on demand). They allow sending and downloading any type of static digital content, not just audio and video (for example: documents, images, etc.). Each request handles one request per call, following a request/response model, and they are not optimized for continuous or real-time transmissions. Unlike WS, microservices are designed to deliver lightweight and specific responses, generally focused on business logic, control, metadata, and content orchestration. Although they can deliver content, they are not intended for transferring large files.

On the other hand, WebRTC and live streaming platforms are geared towards delivering dynamic content, and in most cases, in real time or with very low latency. These methods are used for video calls, live broadcasts, and direct user interaction. Finally, streaming platforms can deliver static or on-demand content, but they do so in a fragmented way. The content is divided into small segments and, in many cases, the browser or player is responsible for downloading and joining those fragments to play them as a continuous stream.

Table 8. Comparison of content delivery methods.

Method	Architecture	Communication	Transmission	Distribution	Protocol
IPTV	Centralized	Bidirectional	Multicast	Live/On-Demand	IP/RTP/UDP/HTTP
WebRTC	Peer to Peer Centralized	Bidirectional Bidirectional	Unicast	Live	RTP
HLS	Centralized	Unidirectional	Multicast	Live/On-Demand	HTTP
Streaming	Centralized	Unidirectional	Unicast	OnDemand	HTTP/HLS/MPEG-DASH/RTP
Microservices	Decentralized	Unidirectional	Unicast	OnDemand	HTTP, SOAP, MQTT, CoAP
Web Services	Centralized	Unidirectional	Unicast	OnDemand	HTTP

5.7 Selected Studies

Regardless of the type of WS used to deliver content, four studies were selected, which are the studies by (Zagarese et al., 2015), (Castiglione et al., 2015), (H. Zhang et al., 2016) and (Xu et al., 2019). The attributes that influence content delivery via WS are analyzed in section 5.9.

In (Zagarese et al., 2015) the authors propose a scalable publish/subscribe architecture that uses SOAP-based WS, and an extension that allows interoperability to attach files and send them through the response to a Web service request. The pub/sub architecture is composed of three patterns, 1 publisher–1 Subscriber, 1 Publisher–N Subscribers, and N Publishers–1 Subscriber; in this architecture, a client subscribes to a specific type of event, according to (i) content, (ii), topic, or (iii) interest; the Event Cloud (EC) manages the subscriptions and handles the events, the publisher launches the publications and notifies the EC, and the subscriber retrieves the notifications asynchronously. The method to deliver files or content is through attachments instead of embedded in the SOAP message, in which a middleware sends a notification with a link to subscribers with the instructions to download the resource.

In (Castiglione et al., 2015) the authors present a cloud architecture to secure manage 3D medical images regardless of the computational and networking capabilities. The images are compressed on the basis of a predictive technique and delivered with a watermark to the clients. In this study the delivery of content is through SOAP-MTOM, which is a method used for large data transmission as part of the SOAP message as attached files instead of embedded in the SOAP message.

In (H. Zhang et al., 2016) the authors propose a cloud platform that employs Docker containers with a video microservices management environment and high service capacity, as well as a fine-grained predictive resource provisioning approach that can periodically predict workload. The predictive approach performs proactive resource provisioning for the video microservices, which uses a service similarity matching technique and the time series nearest neighbor regression method. These allow for efficient prediction of resource requirements and dynamically optimize resource usage based on predictive results to ensure service quality. Although it delivers video content via web services, the delivery method is not specified, nor are any details about the content provided.

In (Xu et al., 2019) the authors propose a prediction method for selecting WS, in which the microservices architecture is used to deliver video content, each microservice delivers a video of approximately one-minute length. The prediction is determined from the Particle Swarm Optimization algorithm and two heuristic methods, explicit factor analysis and linear regression to obtain the optimal service according to the QoS calculation of the transmission, and it is determined from the network load and bandwidth, in addition to the frame rate, video resolution, and data rate. Although the study focuses on a prediction method, the prediction method considers computing network elements, such as network load and bandwidth. In any case, the study not defines the delivery method nor indicate the integration of the content with the web service.
Table 9 presents a comparative analysis of the advantages and disadvantages of the four selected studies with respect to their content delivery capabilities, such as the content delivery method identified as an advantage or the use of multiple messages per notification identified as a disadvantage.

Table 9. Advantages and disadvantages of selected studies.

Study	Advantages	Disadvantages
(Zagarese et al., 2015)	Multiple resources can be submitted. Use SOAP WS-I.	Delay sending messages. Multiples messages per notifications. The resource is not sent with the service, instead a link is sent to facilitate its download
(Castiglione et al., 2015)	The image is compressed and delivered. Use SOAP-MTOM.	The attributes of devices are not considered.
(H. Zhang et al., 2016)	The prediction component computes resource requirements based on historical data and video service data.	Does not indicate how the video service works.
(Xu et al., 2019)	Selection of services according to the characteristics of the video service data. Results are analyzed in real time.	Does not indicate how the video service works.

5.8 Quality Assessment

As mentioned above, the quality of the studies is based on the recognition of the scientific community. However, an assessment was performed according to the protocol of the experiment. 33The assessment result is contrasting, the studies by (Zagarese et al., 2015), (Castiglione et al., 2015) and (H. Zhang et al., 2016), are well structured, readable, and the experiment is well defined; the study by (Xu et al., 2019) has deficiencies, since it does not provide enough information on the video service architecture. In a similarity comparison, four attributes were reviewed: applicability indicates how applicable it is to our study; validity indicates whether the data and results are valid; reliability indicates whether the results and data of the samples are reliable and statistically accurate; relevance indicates the magnitude of the relevance according to the research question with respect to our study. Regarding the mentioned attributes, in (Zagarese et al., 2015) WS are used and deliverability is evaluated. Validity is maintained in Medium because statistical methods are not used, the confidence interval is not defined, and the certainty is not given. The study by (Castiglione et al., 2015) despite being well-prepared, it does not have high values because it does not pursue the purpose of the research question, and in the tests and results no statistical certainty is obtained, nor is a level of confidence given. In the case of (H. Zhang et al., 2016), the applicability and relevance values are medium because is focused on the prediction of results, although it considers server attributes such as memory, CPU, and application containerization, there is no statistical certainty that the differences between the results of the algorithms are reliable. Lastly, in (Xu et al., 2019) the assessment values are low because the focus of the study is different from the research question.

Based on the information of the selected studies, it is unlikely that their results can be replicated. The lack of detailed descriptions constrained experimental setups and restricts replication.
The studies by Xu et al. (2019) and H. Zhang et al. (2016), as previously discussed, they indicate the use of microservices for video content delivery without providing sufficient information about the Content and the WS. Consequently, their experiments cannot be replicated, and their findings cannot be reliably transferred to other environments. Both studies focus primarily on service selection strategies rather than on measurable delivery performance.

Conversely, the studies by Zagarese et al. (2015) and Castiglione et al. (2015) rely on SOAP-based Web Services. The former implements WS-I for file attachment, an approach to binary content transfer, while the latter employs MTOM to deliver 3D medical images. Among these, only the study by Zagarese et al. (2015) evaluates delivery capability under specified conditions. In contrast, the remaining studies emphasize intrinsic attributes of the delivered content rather than performance metrics, further limiting the extent to which their results can be applied to broader or more diverse contexts. The results presented in Table 10 for applicability, reliability, and relevance reflect the replicability of the experiments reported in the selected studies.

Table 10. Quality assessment of the selected studies.

Study	(Zagarese et al., 2015)	(Castiglione et al., 2015)	(H. Zhang et al., 2016)	(Xu et al., 2019)
Context	High	High	High	Medium
Legibility	High	High	High	Medium
Experimentation	High	Medium	High	Medium
Analysis	Medium	High	High	Medium
Applicability	High	Medium	Medium	Low
Validity	Medium	Medium	High	Medium
Reliability	High	Medium	High	Medium
Relevance	High	Medium	Medium	Low

5.9 Deliverability

For the purposes of this work, the term deliverability is defined as the ability to deliver content. In this context, four studies are relevant to answer the RQ. The deliverability evaluation proposed in (Zagarese et al., 2015), considers quality attributes such as event size, set of attachments, latency, and network throughput. In the studies by (Castiglione et al., 2015; H. Zhang et al., 2016; Xu et al., 2019) the deliverability or delivery capability is not evaluated. However, in the content management considering attributes such as network, CPU and server memory, video content characteristics such as frame rate and data rate. Across these studies, SOAP MTOM, server memory, and containerization are the most important characteristics.
Some studies propose attributes such as response time, availability, throughput, latency, etc., and they are attributes that are determined by the capacity of the computing network infrastructure. In this SLR process, a set of important attributes to support content delivery through WS were identified:

- **Server Cache (SC):** Distributed computing technique that allows to store replicated content and ready for consumption, and is geographically located closer to the user. In the studies by (Kilanioti & A. Papadopoulos, 2017; Shi et al., 2020) the authors consider the use of a Server Cache as a key attribute for content delivery.
- **Load Balancing (LB):** A server attribute that may support content delivery. A replication architecture using a load balancer allows a request to be responded successfully than a single architecture or a non-replicated architecture. In (Gundu et al., 2021) the authors analyze a set of algorithms to select the best algorithm in Load Balancing environments, in terms of memory, CPU, bandwidth, response time, processing time, etc., and indicate that a Load Balancing architecture increases the quality of service.
- **Fault Tolerance (FT):** An attribute of the server used in the replication architecture that redirects a response in the case of no response. In the study by (Akpinar et al., 2017) the authors propose a fault-tolerance architecture to deploy critical applications that must be set in a high-availability environment.
- **Communication Protocol (CP):** Some of the studies found in the SLR search consider the analysis of the communication protocol for content delivery. For example, in the study by (Bhattacharyya et al., 2018) the authors propose a solution for content delivery through streaming using the CoAP protocol. Also, in the study by (Sheltami et al., 2018) the authors propose the use of the CoAP protocol to enable efficient real-time streaming in constraint devices. In (Tran et al., 2018) the authors propose a quality model for adaptive streaming of HTTP, based on the concept of a sliding window of video segments over a session. In (Ramadan & Mohamed, 2016) the authors indicate that the HTTP/2 attributes, such as multiplexing and concurrent responses, are more efficient than HTTP/1.
- **Container Technology (CT):** Technology is more efficient than Virtual Machines, according to (H. Zhang et al., 2016) a container produces a more lightweight and more agile virtualization computing resources.
- **Resource Allocation (RA):** A technique that allows the elasticity of servers, whether vertical by increasing resources such as memory, CPU, or storage, or horizontally by scaling the number of server containers, virtual machines, databases, or nodes. In this context, several studies indicate that resource allocation is one of the most used techniques in the Cloud Computing to attend excess demand for applications such as multimedia or mobile services.

6 Research Directions

No study was found that specifies the elements that should be considered for content delivering. Through a set of attributes, the study by (Zagarese et al., 2015) exposes a way to evaluate the content delivery via WS, but does not consider the server elements. Future work on content delivery methods may fall into the following topics:

- **Delivery Forecasting Methods with Algorithms and Artificial Intelligence:** Determine the deliverability results of quality of service, based on the attributes of a quality model or an architecture type.
- **Network Packet Routing:** In research studies on computer network environments, routing is based on obtaining the shortest path when sending a packet through computer networks. Virtualization of Network Functions (VNF) is a current approach that is used to route network packets.
- **Microservices Architecture:** The microservices paradigm focuses on refining the modularity of components and the granularity of the response service.
- **Content Delivery Networks (CDNs):** These architectural designs are based on the geographical distribution of the servers, so they reduce loading times, latency times, bandwidth costs, and increase the availability of content. Some of the attributes found in this domain are the Cache Server, Fault Tolerance, and Load Balancing.

- **Fog and Edge Computing:** They are responsible for distributing services, such as software, infrastructure, or platforms through the computing network to closer customer locations.

6.1 Current Trends in Service Computing

- Microservices is one of the newest forms of SOA and, in accordance with the authors (Aksakalli et al., 2021) they are an architectural style consisting of small services to form complex software applications. Each microservice is responsible for performing a task. These services can communicate with each other using messaging protocols/services or language-independent protocols like REST.
- Artificial intelligence methods applied to forecasting according to the quality attributes that are required to be evaluated. Based on quality attributes, forecasting methods allow multiple tasks such as the selection, composition and recommendation of services, etc. One of the Artificial Intelligence methods is machine learning (ML); some examples of ML are the studies by (Morariu et al., 2020) that implement ML for resource allocation based on electrical energy saving attributes. In the study by (Song, 2021) ML focuses on predictive methods based on user attributes, WS, and Web conditions.
- Media services or entertainment content platforms such as Netflix, Prime Video, Disney+, etc., provide content to their clients through streaming, which, depending on the relevance or novelty of the content, may be requested by more users. The concern of this type of platform is to use or implement several strategies that allow content delivery at the lowest cost, using the minimum possible of computational resources. In addition, there are other content delivery methods, WebRTC is a method of live or real-time video transmission. The studies by (Oh et al., 2015) and (Xhagjika et al., 2017) deliver content in real time through WebRTC, where the origin of the recording must send the information to the server and from the server to the clients or those interested in viewing the content. Broadcasting is another method of video transmission in which all nodes receive content or data transmission. Another example of media service is IPTV, which according to (Yoon, 2015) is a type of two-way broadcast that combines broadcasting, telecommunications, and related customer information.
- Content delivery networks are characterized by the cloud, fog, and edge computing approach, and a set of server components that provide resources, data, or content closer to end users. CDN provides a significant improvement in service architectures, reducing latency, time, bandwidth, and costs.
- Based on the information from the selected studies, they do not state which would be the best method for content delivery through WS, (Zagarese et al., 2015) uses attachments, and (Castiglione et al., 2015) uses SOAP MTOM. Regarding for service quality attributes, parameters, or methods that must be used for content delivery, some attributes are not considered, they neither define which are the best values, for example, Java Heap Space parameters, communication protocol, nor the type of WS (SOAP or REST) must be used for content delivery. In some studies, some attributes are considered, such as response time, throughput, CPU, RAM, or the size of the content, whether images or videos.

7 Threats to Validity

This section describes potential threats to the validity of this SLR, together with the corresponding mitigation strategies.

- **Related Work Sources.** The related work sources consulted in this SLR are ScienceDirect, ACM, IEEE Xplore and SpringerLink. A manual process was carried out, which means that no automated procedures or tools were used, so this SLR is exposed to human error.
- **Search String and Processing**. Synonyms and keywords were defined and tested exhaustively, so search strings and processing do not affect the results.
- **Academic Perspective.** The studies presented in this SLR come from libraries with an academic approach, from conferences, and are for scientific use. No studies were taken from unrecognized sources.
- **False negatives.** The work was processed manually, although the review and selection processes were repeated three times, bias or the possibility of fakes exists. The process of reading the title and abstract was carried out three times, which significantly reduces the risk of bias. In this SLR process, 79 studies were considered and at the end four studies were selected.
- **Selection Automation.** No type of automatic tool was used, the inclusion and exclusion criteria were defined according to the proposed objective, and therefore after summaries readings and subsequently performing full texts reading, no false exclusions were found.
- **False Positives.** By carrying out the process of reviewing and performing the summaries three times and

subsequently performing full text reading, the required corrections were made.

- **Data Extraction Errors.** In the four selected studies, the tests were carried out with synthetic data, which may represent a risk to validity.
- **Defects.** In all SLR stages, defects were corrected; furthermore, original documents were consulted during data synthesis to confirm the accuracy of the extracted data and the assigned categories.

8 Conclusions

This SLR analyzes the studies that focus on content delivery through WS, findings show that there are few works in this approach since authors consider mainly quality attributes related to networks such as latency, response time, throughput, etc. Also, an important issue is resource allocation on the servers. The most recurring topics in the Cloud Computing area are resource allocation and the orchestration, composition, selection, or classification of services. Computer networking topics include the use of CDN, SDN, VNF, and mobile networks. In the IoT category, the topics are very diverse. Other attributes such as maintainability are mentioned in the studies, but no measurement methods were defined. Therefore, studies that precisely define more attributes and indicators or metrics are needed in order to select and employ WS based on best value attributes. No objective studies were found that determine which is the best implementation of WS, REST or SOAP, to deliver content, although on the SOAP side, important attributes such as WS-I attachments and MTOM were found. In addition, the implementation language is considered in very few studies.

Since the intention of this work is to identify quality attributes involved in the process of content delivery through WS, the main conclusion is that attributes such as vertical and horizontal elasticity, cache server, and fault tolerance should be included in the studies since these attributes contribute for content delivery via WS. In addition, there are no metrics or indicators for content delivery other than the work of (Zagarese et al., 2015), in which the evaluation framework considers network traffic, the number of connected users, desired content, delays, failure rates, and a series of assumptions to develop a purchasing strategy for cloud service providers and balance user demand with costs. As such, this SLR suggests researching more approaches for content delivery and including the quality attributes in WS selection processes. The precise definitions of clients and servers, that is, definitions that include descriptions of elements that are part of them, and research of their best values should be conducted. A proposal is to include the study of (Suri et al., 2019) which analyzes the set of communication protocols for the dissemination of resources. It is important to mention that the objective of the SLR was achieved since several quality attributes that may affect the content delivery through WS were identified, being load balancing, fault tolerance, communication protocols and resource allocation, the most named. However, research works do not provide enough information about the best attribute values and the best combination of attributes, they may be selected when infrastructure for content delivery is considered. These are relevant due to the large number of formal proposals from which to select, compose, orchestrate, and classify web services. Also, they may be used in SLA to identify and guarantee the requirements that a WS must meet.

On the other hand, the relationship between attributes must be determined. In the first case of related attributes, such as maintainability and reusability, the more reusable the code, the more maintainable it is. Throughput and response time attributes are two contrasting; that is, the higher the response time, the lower the throughput. In the last example of non-related attributes, success rate and response time attributes are unrelated; they do not correlate with each other. In this context, there are several methods for multi-criteria decision-making for normalize measurement values across different dimensions (time, number of errors, percentage, CPU, severity, etc.). The most known methods are Analytic Hierarchy Process (AHP), Simple Addition Weight (SAW), Multi-Objective Optimization on the basis of Ratio Analysis (Moora), Analytic Network Process (ANP), Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS).

Finally, the size of the content to be delivered is a subject of study in several works since granularity itself is a very important attribute in areas such as Cloud computing, distributed computing and software engineering. When these areas have complete information about the granularity of content WS, domains such as E-Learning, medical and healthcare services, multimedia, streaming, and many others can significantly benefit.

References

Abdelmaboud, A., Jawawi, D. N. A., Ghani, I., Elsafi, A., & Kitchenham, B. (2015). Quality of service approaches in cloud computing: A systematic mapping study. *Journal of Systems and Software, 101*, 159–179. https://doi.org/10.1016/j.jss.2014.12.015

Abdou, W., Darties, B., & Mbarek, N. (2015). An autonomic message dissemination protocol for vehicular ad hoc networks: A density and priority levels aware approach. *Wireless Networks, 21*(3), 1001–1014. https://doi.org/10.1007/s11276-014-0831-x

Abu-Lebdeh, M., Belqasmi, F., & Glitho, R. (2016). An architecture for QoS-enabled mobile video surveillance applications in a 4G EPC and M2M environment. *IEEE Access, 4*, 4082–4093. https://doi.org/10.1109/ACCESS.2016.2592919

Ahammad, I., Khan, M. A. R., & Salehin, Z. U. (2021). QoS performance enhancement policy through combining fog and SDN. *Simulation Modelling Practice and Theory, 109*, 102292. https://doi.org/10.1016/j.simpat.2021.102292

Akhtar, R., Leng, S., Memon, I., Ali, M., & Zhang, L. (2015). Architecture of hybrid mobile social networks for efficient content delivery. *Wireless Personal Communications, 80*(1), 85–96. https://doi.org/10.1007/s11277-014-1996-4

Akpinar, K., Jafariakinabad, F., Hua, K. A., Nakhila, O., Ye, J., & Zou, C. (2017). Fault-tolerant network-server architecture for time-critical web applications. En *2017 IEEE 15th International Conference on Dependable, Autonomic and Secure Computing (DASC)* (pp. 377–384). IEEE. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.79

Al Ridhawi, I., Kotb, Y., & Al Ridhawi, Y. (2017). Workflow-net based service composition using mobile edge nodes. *IEEE Access, 5*, 23719–23735. https://doi.org/10.1109/ACCESS.2017.2766068

Alam, A. F. B., Soltanian, A., Yangui, S., Salahuddin, M. A., Glitho, R., & Elbiaze, H. (2016). A cloud platform-as-a-service for multimedia conferencing service provisioning. En *2016 IEEE Symposium on Computers and Communication (ISCC)* (pp. 289–294). IEEE. https://doi.org/10.1109/ISCC.2016.7543756

Almadani, B. (2015). Drilling data management in petroleum industry based on RTPS. *Procedia Computer Science, 56*, 325–332. https://doi.org/10.1016/j.procs.2015.07.215

Almadani, B., Al Mamun, A., & Khayyat, A. (2015). Real-time QoS-aware vehicle tracking: An experimental and comparative study. *Procedia Computer Science, 56*, 349–356. https://doi.org/10.1016/j.procs.2015.07.218

Al-Turjman, F. (2020). Intelligence and security in big 5G-oriented IoNT: An overview. *Future Generation Computer Systems, 102*, 357–368. https://doi.org/10.1016/j.future.2019.08.009

Apple Inc. (2023). *HTTP live streaming (HLS) authoring specification for Apple devices*. https://developer.apple.com/documentation/http-live-streaming/hls-authoring-specification-for-apple-devices

Asghari, P., Rahmani, A. M., & Javadi, H. H. S. (2019). Internet of Things applications: A systematic review. *Computer Networks, 148*, 241–261. https://doi.org/10.1016/j.comnet.2018.12.008

Aygun, B., Gunel Kilic, B., Arici, N., Cosar, A., & Tuncsiper, B. (2021). Application of binary PSO for public cloud resources allocation system of video on demand (VoD) services. *Applied Soft Computing, 99*, 106870. https://doi.org/10.1016/j.asoc.2020.106870

Bakri, H., Allison, C., Miller, A., & Oliver, I. (2015). HTTP/2 and QUIC for virtual worlds and the 3D web? *Procedia Computer Science, 56*, 242–251. https://doi.org/10.1016/j.procs.2015.07.204

Barakabitze, A. A., Ahmad, A., Mijumbi, R., & Hines, A. (2020). 5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges. *Computer Networks, 167*, 106984. https://doi.org/10.1016/j.comnet.2019.106984

Barba-Jimenez, C., Ramirez-Velarde, R., Tchernykh, A., Rodríguez-Dagnino, R., Nolazco-Flores, J., & Perez-Cazares, R. (2016). Cloud-based video-on-demand service model ensuring quality of service and scalability. *Journal of Network and Computer Applications, 70*, 102–113. https://doi.org/10.1016/j.jnca.2016.05.007

Baxmann, M., Kárpáti, K., & Baráth, Z. (2025). A systematic review of the content and delivery of clinical knowledge in orthodontic postgraduate programs. *BMC Medical Education, 25*(1). https://doi.org/10.1186/s12909-025-07361-x

Bellido, J., Alarcón, R., & Pautasso, C. (2013). Control-flow patterns for decentralized RESTful service composition. *ACM Transactions on the Web, 8*(1), 1–30. https://doi.org/10.1145/2535911

Bhamare, D., Jain, R., Samaka, M., & Erbad, A. (2016). A survey on service function chaining. *Journal of Network and Computer Applications, 75*, 138–155. https://doi.org/10.1016/j.jnca.2016.09.001

Bhardwaj, T., & Sharma, S. C. (2018). Cloud-WBAN: An experimental framework for cloud-enabled wireless body area network with efficient virtual resource utilization. *Sustainable Computing: Informatics and Systems, 20*, 14–33. https://doi.org/10.1016/j.suscom.2018.08.008

Bhattacharyya, A., Agrawal, S., Rath, H. K., & Pal, A. (2018). Improving live-streaming experience for delay-sensitive IoT applications: A RESTful approach. En *2018 IEEE Globecom Workshops (GC Wkshps)* (pp. 1–7). IEEE. https://doi.org/10.1109/GLOCOMW.2018.8644521

Bolettieri, S., Bruno, R., & Mingozzi, E. (2021). Application-aware resource allocation and data management for MEC-assisted IoT service providers. *Journal of Network and Computer Applications, 181*, 103020. https://doi.org/10.1016/j.jnca.2021.103020

Cánovas, A., Taha, M., Lloret, J., & Tomás, J. (2018). Smart resource allocation for improving QoE in IP multimedia subsystems. *Journal of Network and Computer Applications, 104*, 107–116. https://doi.org/10.1016/j.jnca.2017.12.020

Castiglione, A., Pizzolante, R., De Santis, A., Carpentieri, B., Castiglione, A., & Palmieri, F. (2015). Cloud-based adaptive compression and secure management services for 3D healthcare data. *Future Generation Computer Systems, 43–44*, 120–134. https://doi.org/10.1016/j.future.2014.07.001

Castillo, P. A., Bernier, J. L., Arenas, M. G., Merelo, J. J., & Garcia-Sanchez, P. (2011). SOAP vs. REST: Comparing a master–slave GA implementation. *arXiv Preprint*. https://doi.org/10.48550/arXiv.1105.4978

Chand, D., & Ogul, H. (2020). Content-based search in lecture video: A systematic literature review. En *Proceedings of the 3rd International Conference on Information and Computer Technologies (ICICT 2020)* (pp. 169–176). IEEE. https://doi.org/10.1109/ICICT50521.2020.00034

Chaudhuri, A. (2017). Hierarchical support vector regression for QoS prediction of network traffic data. En *Proceedings of the 1st International Conference on Internet of Things and Machine Learning* (pp. 1–6). ACM. https://doi.org/10.1145/3109761.3158386

Cheng, B., & Hancke, G. P. (2015). A service-oriented architecture for wireless video sensor networks: Opportunities and challenges. En *IECON 2015 – 41st Annual Conference of the IEEE Industrial Electronics Society* (pp. 2667–2672). IEEE. https://doi.org/10.1109/IECON.2015.7392504

Cheng, X., Dale, C., & Liu, J. (2007). Understanding the characteristics of Internet short video sharing: YouTube as a case study. *arXiv Preprint*. http://arxiv.org/abs/0707.3670

Curbera, F., Duftler, M., Khalaf, R., Nagy, W., Mukhi, N., & Weerawarana, S. (2002). Unraveling the Web services web: An introduction to SOAP, WSDL, and UDDI. *IEEE Internet Computing, 6*(2), 86–93. https://doi.org/10.1109/4236.991449

Dantas, J., Matos, R., Araujo, J., Oliveira, D., Oliveira, A., & Maciel, P. (2016). Hierarchical model and sensitivity analysis for a cloud-based VoD streaming service. En *2016 IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)* (pp. 10–16). IEEE. https://doi.org/10.1109/DSN-W.2016.23

de Assunção, M. D., Cardonha, C. H., Netto, M. A. S., & Cunha, R. L. F. (2016). Impact of user patience on auto-scaling resource capacity for cloud services. *Future Generation Computer Systems, 55*, 41–50. https://doi.org/10.1016/j.future.2015.09.001

De Renzis, A., Garriga, M., Flores, A., Cechich, A., Mateos, C., & Zunino, A. (2017). A domain-independent readability metric for web service descriptions. *Computer Standards & Interfaces, 50*, 124–141. https://doi.org/10.1016/j.csi.2016.09.005

De Rivera, D. S., Alcarria, R., Martin, D., Sanchez-Picot, A., Bordel, B., & Robles, T. (2016). Distributed query results and IoT data in a publish–subscribe network implementing user notifications. En *IEEE 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA 2016)* (pp. 778–783). IEEE. https://doi.org/10.1109/WAINA.2016.118

Dilley, J., Maggs, B., Parikh, J., Prokop, H., Sitaraman, R., & Weihl, B. (2002). Globally distributed content delivery. *IEEE Internet Computing*.

Enríquez, R. A. H., Cáceres, J. R. R., & Robles, T. de J. Á. (2022). Accessible interactive systems for deaf users in museums: Systematic mapping review. En *2022 International Conference on Inclusive Technologies and Education (CONTIE)* (pp. 1–5). IEEE. https://doi.org/10.1109/CONTIE56301.2022.10004432

Fan, H., Hussain, F. K., Younas, M., & Hussain, O. K. (2015). An integrated personalization framework for SaaS-based cloud services. *Future Generation Computer Systems, 53*, 157–173. https://doi.org/10.1016/j.future.2015.05.011

Faniyi, F., & Bahsoon, R. (2016). A systematic review of service level management in the cloud. *ACM Computing Surveys, 48*(3), 1–27. https://doi.org/10.1145/2843890

Farhad, S. M., Bappi, M. S. I., & Ghosh, A. (2016). Dynamic resource provisioning for video transcoding in IaaS cloud. En *IEEE International Conference on High Performance Computing and Communications (HPCC)* (pp. 380–384). IEEE. https://doi.org/10.1109/HPCC-SmartCity-DSS.2016.0061

Gao, H., Duan, Y., Shao, L., & Sun, X. (2021). Transformation-based processing of typed resources for multimedia sources in the IoT environment. *Wireless Networks, 27*(5), 3377–3393. https://doi.org/10.1007/s11276-019-02200-6

Gaur, A. S., Budakoti, J., Lung, C.-H., & Redmond, A. (2017). IoT-equipped UAV communications with seamless vertical handover. En *IEEE Conference on Dependable and Secure Computing* (pp. 459–465). IEEE. https://doi.org/10.1109/DESEC.2017.8073829

Ge, H., Jin, Y., & Yang, T. (2016). A QoE-based control system for bitrate adaptation of wireless-accessed video streams. En *4th International Conference on Cloud Computing and Intelligence Systems (CCIS)* (pp. 125–130). IEEE. https://doi.org/10.1109/CCIS.2016.7790238

Gkamas, V., Koutoumanos, A., Alexandris, K., Megalou, E., Paraskevas, M., & Kaklamanis, C. (2016). Integrating the Kaltura video platform with the Photodentro video repository: A case study. En *19th IEEE International Conference on Computational Science and Engineering* (pp. 173–176). IEEE. https://doi.org/10.1109/CSE-EUC-DCABES.2016.180

Grigorik, I. (2013). Making the web faster with HTTP 2.0. *Communications of the ACM, 56*(12), 42–49. https://doi.org/10.1145/2534706.2534721

Gundu, S. R., Panem, C. A., Thimmapuram, A., & Gad, R. S. (2021). Emerging computational challenges in cloud computing and RTEAH algorithm-based solution. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-021-03380-w

Gutierrez-Garcia, J. O., & Sim, K. M. (2015). Agent-based cloud bag-of-tasks execution. *Journal of Systems and Software, 104*, 17–31. https://doi.org/10.1016/j.jss.2015.02.039

Halili, F., & Ramadani, E. (2018). Web services: A comparison of SOAP and REST services. *Modern Applied Science, 12*(3), 175–186. https://doi.org/10.5539/mas.v12n3p175

Hani, A. F. M., Paputungan, I. V., & Hassan, M. F. (2015). Renegotiation in service level agreement management for a cloud-based system. *ACM Computing Surveys, 47*(3), 1–21. https://doi.org/10.1145/2716319

Heller, M. (1998). Introduction: Downloadable audio and video network update—Internet-delivered audio and video.

Holowczak, J., & Houmansadr, A. (2015). CacheBrowser: Bypassing Chinese censorship without proxies using cached content. En *Proceedings of the ACM Conference on Computer and Communications Security (CCS 2015)* (pp. 70–83). ACM. https://doi.org/10.1145/2810103.2813696

Hossain, S. M., Hnat, T., Saleheen, N., Nasrin, N. J., Noor, J., Ho, B. J., Condie, T., Srivastava, M., & Kumar, S. (2017). mCerebrum: A mobile sensing software platform for development and validation of digital biomarkers and interventions. En *Proceedings of the 15th ACM Conference on Embedded Networked Sensor Systems (SenSys 2017)*. ACM. https://doi.org/10.1145/3131672.3131694

Huang, Y., Huang, J., Wu, B., Yao, T., He, S., & Chen, J. (2016). An ontology-based semantic service markup for content-centric networking. En *IEEE International Conference on Services Computing (SCC)* (pp. 794–797). IEEE. https://doi.org/10.1109/SCC.2016.109

Huf, A., Salvadori, I., & Siqueira, F. (2016). A service-oriented approach for integrating broadcast facilities. En *IEEE International Conference on Services Computing (SCC)* (pp. 705–712). IEEE. https://doi.org/10.1109/SCC.2016.97

IBM Corporation. (2018). *WS-I attachments profile*. http://www.ws-i.org/Profiles/AttachmentsProfile-1.0-2004-08-24.html

Jaiganesh, M., Ramadoss, B., Kumar, A. V. A., & Mercy, S. (2015). Performance evaluation of cloud services with profit optimization. *Procedia Computer Science, 54*, 24–30. https://doi.org/10.1016/j.procs.2015.06.003

Jyoti, A., Shrimali, M., Tiwari, S., & Singh, H. P. (2020). Cloud computing using load balancing and service broker policy for IT service: A taxonomy and survey. *Journal of Ambient Intelligence and Humanized Computing, 11*(11), 4785–4814. https://doi.org/10.1007/s12652-020-01747-z

Kamali, M. E., Angelini, L., Caon, M., Carrino, F., Rocke, C., Guye, S., Rizzo, G., Mastropietro, A., Sykora, M., Elayan, S., Kniestedt, I., Ziylan, C., Lettieri, E., Khaled, O. A., & Mugellini, E. (2020). Virtual coaches for older adults' wellbeing: A systematic review. *IEEE Access, 8*, 101884–101902.

Karabey Aksakalli, I., Çelik, T., Can, A. B., & Tekinerdogan, B. (2021). Deployment and communication patterns in microservice architectures: A systematic literature review. *Journal of Systems and Software, 180*, 111014. https://doi.org/10.1016/j.jss.2021.111014

Ketankumar, D. C., Verma, G., & Chandrasekaran, K. (2015). A green mechanism design approach to automate resource procurement in cloud. *Procedia Computer Science, 54*, 108–117. https://doi.org/10.1016/j.procs.2015.06.013

Khan, S., Hussain, F. K., & Hussain, O. K. (2021). Guaranteeing end-to-end QoS provisioning in SOA-based SDN architecture: A survey and open issues. *Future Generation Computer Systems, 119*, 176–187. https://doi.org/10.1016/j.future.2021.02.011

Kilanioti, I., & Papadopoulos, G. A. (2017). Content delivery simulations supported by social network-awareness. *Simulation Modelling Practice and Theory, 76*, 47–66. https://doi.org/10.1016/j.simpat.2017.01.001

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (EBSE Technical Report, Version 2.3). EBSE.

Kumari, S., & Rath, S. K. (2015). Performance comparison of SOAP and REST based web services for enterprise application integration. En *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1656–1660). IEEE. https://doi.org/10.1109/ICACCI.2015.7275851

Le Duc, T., Leiva, R. G., Casari, P., & Östberg, P. O. (2019). Machine learning methods for reliable resource provisioning in edge-cloud computing: A survey. *ACM Computing Surveys, 52*(5), Article 94. https://doi.org/10.1145/3341145

Martin, G., Mendelsohn, N., Nottingham, M., & Ruellan, H. (2005). *XML-binary optimized packaging (XOP)*. World Wide Web Consortium (W3C). https://www.w3.org/TR/xop10/

Li, J., Zheng, X., Watanabe, I., & Ochiai, Y. (2024). A systematic review of digital transformation technologies in museum exhibition. *Computers in Human Behavior, 161*, 108407. https://doi.org/10.1016/j.chb.2024.108407

McCarthy, J., Chaudhry, S. R., Kuppuudaiyar, P., Loomba, R., & Clarke, S. (2021). QoSA-ICN: An information-centric approach to QoS in vehicular environments. *Vehicular Communications, 30*, 100351. https://doi.org/10.1016/j.vehcom.2021.100351

Montella, R., Di Luccio, D., Marcellino, L., Galletti, A., Kosta, S., Brizius, A., & Foster, I. (2017). Processing of crowd-sourced data from an Internet of Floating Things. En *Proceedings of the 12th Workshop on Workflows in Support of Large-Scale Science (WORKS 2017)* (pp. 1–8). ACM. https://doi.org/10.1145/3150994.3150997

Moon, S., Yoo, J., & Kim, S. (2016). Adaptive interface selection over cloud-based split-layer video streaming via multi-wireless networks. *Future Generation Computer Systems, 56*, 664–674. https://doi.org/10.1016/j.future.2015.09.022

Morariu, C., Morariu, O., Răileanu, S., & Borangiu, T. (2020). Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. *Computers in Industry, 120*, 103244. https://doi.org/10.1016/j.compind.2020.103244

Mustafa, S., Nazir, B., Hayat, A., Khan, A. U. R., & Madani, S. A. (2015). Resource management in cloud computing: Taxonomy, prospects, and challenges. *Computers & Electrical Engineering, 47*, 186–203. https://doi.org/10.1016/j.compeleceng.2015.07.021

Oh, H., Ahn, S., Choi, J., & Yang, J. (2015). WebRTC-based remote collaborative online learning platform. En *Proceedings of the 1st Workshop on All-Web Real-Time Systems (AweS 2015)*. ACM. https://doi.org/10.1145/2749215.2749222

Palumbo, F., Aceto, G., Botta, A., Ciuonzo, D., Persico, V., & Pescapé, A. (2021). Characterization and analysis of cloud-to-user latency: The case of Azure and AWS. *Computer Networks, 184*, 107693. https://doi.org/10.1016/j.comnet.2020.107693

Pautasso, C., Zimmermann, O., & Leymann, F. (2008). RESTful web services vs. "big" web services. En *Proceedings of the 17th International Conference on World Wide Web (WWW 2008)* (pp. 805–814). ACM. https://doi.org/10.1145/1367497.1367606

Quadri, C., Gaito, S., Bruschi, R., Davoli, F., & Rossi, G. P. (2018). A MEC approach to improve QoE of video delivery service in urban spaces. En *2018 IEEE International Conference on Smart Computing (SMARTCOMP)* (pp. 25–32). IEEE. https://doi.org/10.1109/SMARTCOMP.2018.00095

Ramadan, N., & Mohamed, I. (2016). Impact of implementing HTTP/2 in web services. *International Journal of Computer Applications, 147*(9), 27–32. https://doi.org/10.5120/ijca2016911182

Rasbi, M. A. A., & Singh, A. V. (2017). Need and scope of private cloud technology for public authority for radio & television in Oman. En *2017 6th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)* (pp. 523–528). IEEE. https://doi.org/10.1109/ICRITO.2017.8342484

Rizvi, S., Roddy, H., Gualdoni, J., & Myzyri, I. (2017). Three-step approach to QoS maintenance in cloud computing using a third-party auditor. *Procedia Computer Science, 114*, 83–92. https://doi.org/10.1016/j.procs.2017.09.014

Ruan, Z., & Ye, X. (2019). Cost-optimized video dissemination over heterogeneous cloud with SLAs support. *IEEE Access, 7*, 42874–42888. https://doi.org/10.1109/ACCESS.2019.2908176

Said, O., Albagory, Y., Nofal, M., & Al Raddady, F. (2017). IoT-RTP and IoT-RTCP: Adaptive protocols for multimedia transmission over Internet of Things environments. *IEEE Access, 5*, 16757–16773. https://doi.org/10.1109/ACCESS.2017.2726902

Roberts, N. J., Kidd, L., Kirkwood, K., Cross, J., & Partridge, M. R. (2018). A systematic review of the content and delivery of education in pulmonary rehabilitation programmes. *Respiratory Medicine, 145*, 161–181. https://doi.org/10.1016/j.rmed.2018.11.002

Sakthidasan, K., Gao, X.-Z., Devabalaji, K. R., & Mohana Roopa, Y. (2021). Energy-based random repeat trust computation approach and reliable fuzzy and heuristic ant colony mechanism for improving QoS in WSN. *Energy Reports, 7*, 7967–7976. https://doi.org/10.1016/j.egyr.2021.08.121

Santos, H., Alencar, D., Meneguette, R., Rosário, D., Nobre, J., Both, C., Cerqueira, E., & Braun, T. (2020). A multi-tier fog content orchestrator mechanism with quality of experience support. *Computer Networks, 177*, 107288. https://doi.org/10.1016/j.comnet.2020.107288

Sarker, S. A., Rahman, M., Muslim, N., & Islam, S. (2018). Performance analysis of video streaming at the edge and core cloud. En *5th International Conference on Networking, Systems and Security (NSysS 2018)* (pp. 1–7). IEEE. https://doi.org/10.1109/NSysS.2018.8631370

Savolainen, S., & Rinne, I. (2015). Visualising spatial web service growth across Europe. En *Proceedings of the 19th International Academic Mindtrek Conference* (pp. 191–193). ACM. https://doi.org/10.1145/2818187.2818274

Seraoui, Y., Bellafkih, M., & Raouyane, B. (2016). A high-performance and scalable distributed storage and computing system for IMS services. En *International Conference on Cloud Computing Technologies and Applications (CloudTech 2016)* (pp. 335–342). IEEE. https://doi.org/10.1109/CloudTech.2016.7847718

Sheltami, T. R., Shahra, E. Q., & Shakshuki, E. M. (2018). Fog computing: Data streaming services for mobile end-users. *Procedia Computer Science, 134*, 289–296. https://doi.org/10.1016/j.procs.2018.07.173

Shen, J., & Li, H. (2021). The "real images" of engineering ethics education: An analysis based on systematic literature review (SLR). En *IEEE Global Engineering Education Conference (EDUCON 2021)* (pp. 719–725). IEEE. https://doi.org/10.1109/EDUCON46332.2021.9454004

Shi, L., Wang, X., & Ma, R. T. B. (2020). On multi-resource procurement in Internet access markets: Optimal strategies and market equilibrium. *Performance Evaluation, 143*, 102139. https://doi.org/10.1016/j.peva.2020.102139

Silva, T., Kamienski, C., Fernandes, S., & Sadok, D. (2015). A flexible DHT-based directory service for information management. *Peer-to-Peer Networking and Applications, 8*(3), 512–531. https://doi.org/10.1007/s12083-014-0277-z

Singh, S., Chana, I., & Buyya, R. (2020). Agri-Info: Cloud-based autonomic system for delivering agriculture as a service. *Internet of Things, 9*, 100131. https://doi.org/10.1016/j.iot.2019.100131

Stockhammer, T. (2011). Dynamic adaptive streaming over HTTP (DASH): Standards and design principles. En *Proceedings of the Second Annual ACM Conference on Multimedia Systems* (pp. 133–144). ACM. https://doi.org/10.1145/1943552.1943572

Soltanian, A., Belqasmi, F., Yangui, S., Salahuddin, M. A., Glitho, R., & Elbiaze, H. (2018). A cloud-based architecture for multimedia conferencing service provisioning. *IEEE Access, 6*, 9792–9806. https://doi.org/10.1109/ACCESS.2018.2794258

Song, Y. (2021). Web service reliability prediction based on machine learning. *Computer Standards & Interfaces, 73*, 103466. https://doi.org/10.1016/j.csi.2020.103466

Soni, A., & Ranga, V. (2019). API features individualizing of web services: REST and SOAP. *International Journal of Innovative Technology and Exploring Engineering, 8*(9 Special Issue), 664–671. https://doi.org/10.35940/ijitee.I1107.0789S19

Su, G.-M., Su, X., Bai, Y., Wang, M., Vasilakos, A. V., & Wang, H. (2016). QoE in video streaming over wireless networks: Perspectives and research challenges. *Wireless Networks, 22*(5), 1571–1593. https://doi.org/10.1007/s11276-015-1028-7

Suri, N., Breedy, M. R., Marcus, K. M., Fronteddu, R., Cramer, E., Morelli, A., Campioni, L., Provosty, M., Enders, C., Tortonesi, M., & Nilsson, J. (2019). Experimental evaluation of group communications protocols for data dissemination at the tactical edge. En *2019 International Conference on Military Communications and Information Systems (ICMCIS)* (pp. 1–8). IEEE. https://doi.org/10.1109/ICMCIS.2019.8842801

Tellez, N., Jimeno, M., Salazar, A., & Niño-Ruiz, E. D. (2019). Container-based architecture for optimal face-recognition tasks in edge computing. En *Proceedings of the 4th ACM/IEEE Symposium on Edge Computing* (pp. 301–303). ACM. https://doi.org/10.1145/3318216.3363323

Tian, C., Wang, Y., Luo, Y., Jiang, H., Liu, W., Wu, J., & Yin, H. (2016). Minimizing content reorganization and tolerating imperfect workload prediction for cloud-based video-on-demand services. *IEEE Transactions on Services Computing, 9*(6), 926–939. https://doi.org/10.1109/TSC.2015.2416733

Tihomirovs, J., & Grabis, J. (2016). Comparison of SOAP and REST based web services using software evaluation metrics. *Information Technology and Management Science, 19*(1), 92–97. https://doi.org/10.1515/itms-2016-0017

Tran, H. T. T., Ngoc, N. P., Hossfeld, T., & Thang, T. C. (2018). A cumulative quality model for HTTP adaptive streaming. En *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)* (pp. 1–6). IEEE. https://doi.org/10.1109/QoMEX.2018.8463414

Tudoran, R., Costan, A., Nano, O., Santos, I., Soncu, H., & Antoniu, G. (2016). JetStream: Enabling high throughput live event streaming on multi-site clouds. *Future Generation Computer Systems, 54*, 274–291. https://doi.org/10.1016/j.future.2015.01.016

World Wide Web Consortium (W3C). (2023). *WebRTC: Real-time communication in browsers*. https://www.w3.org/TR/webrtc/

Wamser, F., Seufert, M., Höfner, S., & Tran-Gia, P. (2016). Concept for client-initiated selection of cloud instances for improving QoE of distributed cloud services. En *Proceedings of the 2016 Workshop on QoE-Based Analysis and Management of Data Communication Networks* (pp. 49–54). ACM. https://doi.org/10.1145/2940136.2940143

Wang, C., Kim, H., & Morla, R. (2017). QWatch: Detecting and locating QoE anomaly for VoD in the cloud. En *IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2016)* (pp. 126–133). IEEE. https://doi.org/10.1109/CloudCom.2016.0034

Wen, Y., Hu, H., & Liu, F. (2017). Embracing social big data in wireless system design. *Journal of Communications and Information Networks, 2*(1), 81–96. https://doi.org/10.1007/s41650-017-0007-9

Wu, Y., Wu, C., Li, B., Zhang, L., Li, Z., & Lau, F. C. M. (2015). Scaling social media applications into geo-distributed clouds. *IEEE/ACM Transactions on Networking, 23*(3), 689–702. https://doi.org/10.1109/TNET.2014.2308254

Wu, Z., Lu, Z., Hung, P. C. K., Huang, S.-C., Tong, Y., & Wang, Z. (2019). QaMeC: A QoS-driven IoVs application optimizing deployment scheme in multimedia edge clouds. *Future Generation Computer Systems, 92*, 17–28. https://doi.org/10.1016/j.future.2018.09.032

Xavier, G. P., & Kantarci, B. (2018). A survey on the communication and network enablers for cloud-based services: State of the art, challenges, and opportunities. *Annals of Telecommunications, 73*(3–4), 169–192. https://doi.org/10.1007/s12243-018-0629-4

Xhagjika, V., Escoda, O. D., Navarro, L., & Vlassov, V. (2017). Load and video performance patterns of a cloud-based WebRTC architecture. En *2017 IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)* (pp. 739–744). IEEE. https://doi.org/10.1109/CCGRID.2017.118

Xu, Z., Zhang, H., & Huang, H. (2019). A performance-aware selection strategy for cloud-based video services with micro-service architecture. En *Proceedings of the ACM Multimedia Asia* (pp. 1–6). ACM. https://doi.org/10.1145/3338533.3366609

Yamada, J., Blum, N., Carella, G., Kanamaru, N., Uchida, N., & Magedanz, T. (2015). A platform for converged, feature-based real-time communications. En *2015 18th International Conference on Intelligence in Next Generation Networks (ICIN)* (pp. 200–207). IEEE. https://doi.org/10.1109/ICIN.2015.7073832

Yoon, C. (2015). Open IPTV convergence service creation and management using service delivery platform. En *2015 17th International Conference on Advanced Communication Technology (ICACT)* (pp. 607–613). IEEE. https://doi.org/10.1109/ICACT.2015.7224930

Zagarese, Q., Canfora, G., Zimeo, E., Alshabani, I., Pellegrino, L., Alshabani, A., & Baude, F. (2015). Improving data-intensive EDA performance with annotation-driven laziness. *Science of Computer Programming, 97*(Part 2), 266–279. https://doi.org/10.1016/j.scico.2014.03.007

Zaher, M., Alawadi, A. H., & Molnár, S. (2021). Sieve: A flow scheduling framework in SDN-based data center networks. *Computer Communications, 171*, 99–111. https://doi.org/10.1016/j.comcom.2021.02.013

Zamani, A. R., Balouek-Thomert, D., Villalobos, J. J., Rodero, I., & Parashar, M. (2020). An edge-aware autonomic runtime for data streaming and in-transit processing. *Future Generation Computer Systems, 110*, 107–118. https://doi.org/10.1016/j.future.2020.03.037

Zhang, H., Ma, H., Fu, G., Yang, X., Jiang, Z., & Gao, Y. (2016). Container-based video surveillance cloud service with fine-grained resource provisioning. En *2016 IEEE 9th International Conference on Cloud Computing (CLOUD)* (pp. 758–765). IEEE. https://doi.org/10.1109/CLOUD.2016.0105

Zhang, Z., Jiang, X., & Xi, H. (2016). Joint resource allocation and traffic management for cloud video distribution over software-defined networks. En *2016 8th IEEE International Conference on Communication Software and Networks (ICCSN)* (pp. 407–411). IEEE. https://doi.org/10.1109/ICCSN.2016.7586692

Zhao, S., & Medhi, D. (2017). SDN-assisted adaptive streaming framework for tile-based immersive content using MPEG-DASH. En *2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)* (pp. 1–6). IEEE. https://doi.org/10.1109/NFV-SDN.2017.8169831

Zou, L., Trestian, R., & Muntean, G. M. (2015). E2DOAS: User experience meets energy saving for multi-device adaptive video delivery. En *IEEE INFOCOM Workshops 2015* (pp. 444–449). IEEE. https://doi.org/10.1109/INFCOMW.2015.7179425