Challenges of QoE Management for Cloud Applications

Tobias Hoßfeld, University of Würzburg
Raimund Schatz, Telecommunications Research Center Vienna — FTW
Martin Varela, VTT Technical Research Centre of Finland
Christian Timmerer, Alpen-Adria-Universität Klagenfurt

This work was supported in part by the EC in the context of the QUALINET (COST IC 1003) and ALICANTE (FP7-ICT-248652) projects.

Abstract

Cloud computing is currently gaining enormous momentum due to a number of promised benefits: ease of use in terms of deployment, administration, and maintenance, along with high scalability and flexibility to create new services. However, as more personal and business applications migrate to the cloud, service quality will become an important differentiator between providers. In particular, quality of experience as perceived by users has the potential to become the guiding paradigm for managing quality in the cloud. In this article, we discuss technical challenges emerging from shifting services to the cloud, as well as how this shift impacts QoE and QoE management. Thereby, a particular focus is on multimedia cloud applications. Together with a novel QoE-based classification scheme of cloud applications, these challenges drive the research agenda on QoE management for cloud applications.

Introduction

Clouds in the Earth's atmosphere have been studied by meteorologists already for centuries. In contrast, clouds in the information and communication technology (ICT) domain are a very recent phenomenon, albeit one that enjoys spectacularly growing interest from academic and industry stakeholders alike. The main reason behind this development is that cloud computing promises significant economic advantages to ICT users (mostly in terms of cost savings) by moving services, computation, and data to location-transparent centralized facilities or providers. This way, information assets can be more easily accessed and shared with significantly greater flexibility and scalability. Furthermore, moving resources into the cloud facilitates data sharing and multi-user collaboration, which enables novel application concepts and features like cloud gaming and automated tagging of photos.

For these reasons, cloud services are becoming more and more pervasive, not only throughout enterprises aiming to outsource parts of their IT activities to third-party data centers or platforms like Salesforce.com. An increasing number of cloud-based services like DropBox, YouTube, and Google Mail are also finding their way to the desktops of private consumers.

However, the growing presence of cloud-based services creates new problems for both users and providers, resulting in a number of challenges that need to be addressed in order to ensure successful adoption of this new paradigm. Besides well-known and frequently raised issues of privacy and security, a major problem the cloud computing ecosystem faces is the ongoing commoditization of cloud services caused by intensive cost-driven competition among providers: customers expecting significant economic benefits, reduced deployment costs, and the perfect competition characteristics of remote computing exert strong pricing pressures that have already led to near-zero price points in today's cloud markets [1]. In such an environment, providers have to consider another differentiator than just price: the quality of their service. If performance levels do not reach expectations (or become unpredictable) because quality is compromised too much, customers will reject the service or refuse adoption. On the other hand, meeting or exceeding expectations enhances a cloud provider's reputation and increases levels of utilization and adoption. Nonetheless, service providers still have to invest economically in order to remain in business; thus, any service quality improvement also has to be actually perceived and valued by the end customer or end user in order to make a difference. Ensuring that the users' experience remains at least as good as it was before adoption therefore will be one of the key factors in promoting new cloud services and migrations of existing ones.

For these reasons, the concept of quality of experience (QoE) has the potential to become the guiding paradigm for managing quality in the cloud. Being linked very closely to the subjective perception of the end user, QoE enables a broader, more holistic understanding of the factors that influence the performance of systems, and thus complements traditional, technology-centric concepts such as quality of service (QoS). This strictly user-centric focus of QoE is also reflected in its most widespread definition as "overall acceptability of an application or service, as perceived subjectively by the end user" [2]. Consequently, understanding and managing QoE of cloud services requires a multidisciplinary view that integrates technology, user, and business aspects of the end-user quality assessment.

This article introduces the concept of QoE to understanding and managing the quality of cloud services, and provides concrete instances of its use in the media services domain. By taking on a strictly user-centric perspective, it discusses the challenges that arise from the adoption of this new technology and how they affect the quality of typical cloud applications.

The remainder of this article is structured as follows. We provide an overview of typical cloud applications along with commonly used taxonomies for classifying them. Additionally, a new service classification scheme is introduced that is better aligned with the end-user perspective and thus QoE. We outline the most relevant QoE challenges and research questions when shifting services to the cloud. We discuss the resulting challenges for managing the QoE of cloud services along with promising approaches suggested by current research. Finally, we present our conclusions together with an outlook on future research required to bring the full potential of QoE management to the cloud.

Classification of Cloud Applications

Similar to physical clouds, cloud computing can take many different sizes and shapes, particularly when viewed from different perspectives. This section provides an overview of typical cloud applications by discussing various categorization schemes commonly used for cloud services. Finally, it introduces a new classification scheme aligned to the end-user experience and usage domain that will serve as the guiding framework for discussing QoE requirements and management challenges in this article.

Traditional Classification Schemes

Most commonly, cloud computing services are divided into three areas according to the *service delivery model* used [3] and compose the cloud computing stack (from bottom to top):

Infrastructure as a service (IaaS): It offers raw network, computation, and storage infrastructure on which customers can deploy and run arbitrary software (e.g., Amazon's EC2 and GoGrid).

Platform as a service (PaaS): It provides an application environment that enables development and deployment of applications without having to take care of hosting them; for example, Google's AppEngine and Salesforce's Force.com.

Software as a service (SaaS): It uses a provider's applications running on a cloud infrastructure, allowing for limited configurations by the user, such as customer relationship management (CRM) applications like Salesforce.com or office applications like GoogleDocs.

When viewed from a more organizational perspective, cloud services can be deployed in the following four ways [4]:

Public cloud: In this model, several customers access cloud services hosted by the cloud vendor at the vendor's premises. Resources are shared between all clients who have no control over where the underlying infrastructure and services are hosted.

Private cloud: In this type of implementation, services or computing resources are dedicated to a particular organization and not shared with other organizations. The resources are either hosted on premises or off-premises by a third party provider.

Hybrid cloud: In this scenario a client uses private and public clouds, or any other combination of deployment models. A typical use case is the on-demand usage of additional public cloud infrastructure for peak load situations.

Classification Scheme from the End-User Perspective

The two classification schemes above perfectly describe different cloud application models from a technical or organizational deployment point of view. However, when discussing QoE management for the cloud, the end-user perspective has to be taken into account. In this context, end user means the person using a cloud application or consuming a cloud service. The goal of QoE management is then to deliver the cloud application to the end user at high quality, at best while minimizing the costs of the different players of the cloud computing stack (IaaS, PaaS, SaaS) and the underlying network providers: telecommunications companies (telcos) and Internet service providers (ISPs). The QoE requirements and resulting technical requirements from cloud applications differ in the usage dimensions including the degree of multimedia intensity, interactivity, primary usage domain, and service complexity.

Multimedia intensity describes to which extent the cloud application:
- Focuses on multimedia contents
- Allows the end user navigational control of the content progress or manipulation thereof

Degree of interactivity covers:
- Interactions between the end user and the cloud (human-computer interaction), for exammple, to send user commands
- Interactions between different end users (human-to-human communication), such as chat

Primary usage domain separates the major usage and distinguishes between:
- Business applications like office products
- Personal applications like video streaming

Service complexity aggregates the complexity in terms of:

The QoE requirements and resulting technical requirements from cloud applications differ in the usage dimensions including the degree of multimedia intensity, interactivity, primary usage domain, and service complexity.

A large number of technical factors influence the quality of Cloud services and applications as perceived by the end user. These technical factors differ for each individual cloud application and cannot be mapped directly by the introduced classification scheme.

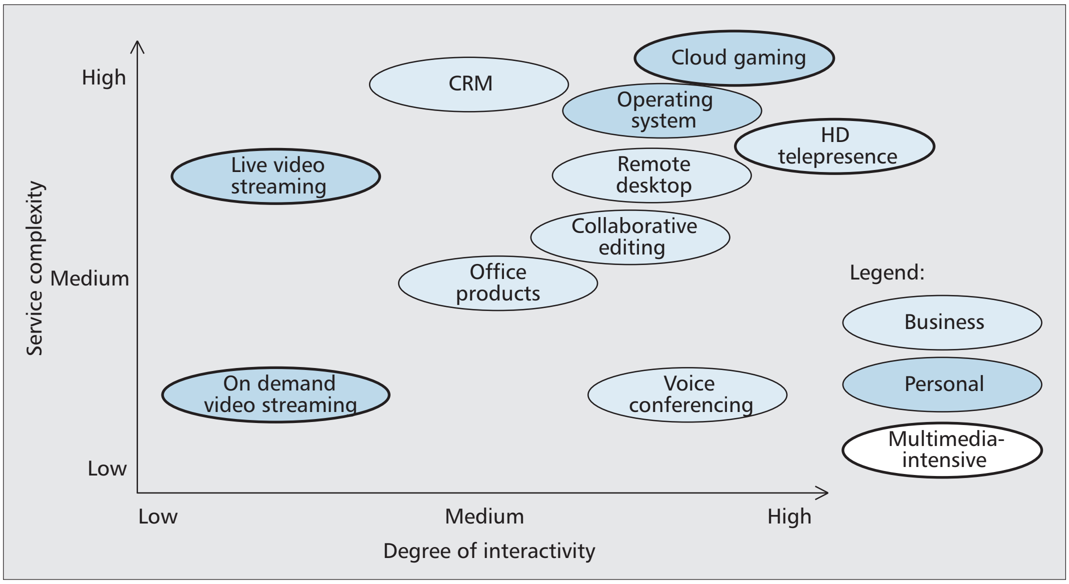

Figure 1. *QoE-based classification of different Cloud applications with regard to level of interactivity, service complexity, usage domain and multimedia-intensity.*

- Technical realization of a cloud and its physical and logical distribution
- Service provisioning and management during business operation, which also includes machine-to-machine communication among various entities of the cloud

Figure 1 depicts the classification scheme for cloud applications from an end-user perspective.[1] The applications are categorized along the above introduced usage dimensions, which directly influence its QoE requirements. For example, business applications are used in corporate environments and are often mission critical, which considerably decreases tolerance for errors and outages.

A large number of technical factors influence the quality of cloud services and applications as perceived by the end user. These technical factors differ for each individual cloud application and cannot be mapped directly by the above introduced classification scheme. However, the key performance factors on cloud QoE have to be monitored and controlled for proper QoE management later on. This includes IaaS performance factors for:

- Networking, like latency and bandwidth
- Computing, like CPU consumption
- Storage, in terms of data volumes

Additionally, technical PaaS/SaaS parameters like database qualities affect the performance of the platform and hence QoE. Regarding SaaS, QoE is also influenced by the user base characteristics in terms of number of users, but also the demographics of the user (considering day-night usage patterns, determining the actual traffic load, etc.).

QoE-based Classification of Cloud Applications

Typical cloud applications are next classified according to the introduced QoE-based scheme. Located on the upper right of Fig. 1, cloud gaming is a highly interactive, multimedia-intensive entertainment service consumed predominantly in private residential contexts. Each player's audio-visual content is rendered on the server side according to his/her game play. Thus, multimedia content has to be created and transmitted to each client at low latency and high synchronicity in order to ensure a good gaming experience. For these reasons, cloud gaming can be considered a fairly complex service, from both the technological and end-user points of view. A less complex application that is almost exclusively used in business contexts is high-definition (HD) telepresence, that is, real-time conferencing for conducting meetings between geographically dispersed participants. HD telepresence services rely on high-bandwidth connections, real-time audio-visual processing, and dedicated hardware such as large screens to support high levels of immersion and natural interaction between participants as known from face-to-face meetings.

Other common cloud services for enterprises are outsourced CRM solutions (e.g., Salesforce.com) and cloud-hosted office suites (e.g., Microsoft Office 365, Google Docs, LotusLive). These are typically delivered as web applications in a SaaS fashion to the end user who accesses them through a web browser. A key challenge in this context is to provide the user experience known from existing desktop applications in a web application environment.

All application categories discussed so far have in common that they feature medium to high levels of interactivity, typically resulting in QoE requirements like low latency and high responsiveness. In contrast, online video streaming services like YouTube and Hulu are less interactive, since their primary use case is the consumption of audio-visual content. Due to the high data volumes and bandwidths involved, cloud technologies are not only used for content storage but also as content delivery networks (CDNs). In this respect, CDNs increase availability, bandwidth utilization, and speed by caching and replicating content over multiple

[1] *Note that this classification scheme is also applicable for non–cloud applications, for example, realized by client–server or peer-to-peer (P2P) technology.*

geographically disperse sites. Regarding service complexity, live video streaming services (e.g., for coverage of sports events) are more challenging in terms of QoE than on-demand services, since content has to be delivered in real time at minimum latency. These features require high service complexity and distributed processing power to render and deliver personalized video streams to the end user in real time.

Due to varying usage contexts, user requirements, and underlying technologies, the above application categories differ in terms of QoE and the QoE management challenges they pose to cloud service providers.

QoE Challenges in the Cloud

Migrating existing services to the cloud, or creating entirely new cloud-based services, poses a new set of technical challenges that cannot be ignored when considering the quality perceived by the service's users. In some cases, technical issues can be simplified by moving to the cloud; for example, scalability is usually much simpler to achieve and manage in a cloud environment. For the most part, however, moving a service to the cloud introduces new technical challenges that do not exist, or at least are not as pronounced, in a more traditional environment. Examples of such issues are:

- Artifacts introduced due to (increase of) the network (distance) between the user and the service
- Resource management problems due to collocation
- Having multiple parties involved in providing the service where before there were one or two
- Geographical distribution of the user base

QoE Challenges for Multimedia Cloud Applications

Depending on the type of service considered, some of these factors might have a significant impact on the quality experienced by its users. For instance, the performance issues of a cloud-based gaming service such as OnLive are not at all the same as playing a game on a local PC [5]. In the case of cloud-based gaming, the network plays a critical role in determining the users' experience, as very strict requirements on latency and bandwidth need to be met. For example, in cloud-based multiplayer online games the game is rendered on the server and streamed to the client. Hence, the player no longer depends on a specific gaming hardware and is able to use common consumer electronics as long as a broadband Internet connection is available and the ability to display HD video. While this may reduce hardware costs for users, it also raises new challenges for service quality in terms of bandwidth and latency for the underlying network. For cloud gaming, no QoE models are available for mapping the network QoS to QoE.

Although multimedia QoE has been studied for a long time, there are issues for which no answers are yet available, especially when considered in a cloud setting. For example, cloud-based video streaming services such as YouTube mostly rely on HTTP/TCP, resulting in completely different behavior in terms of quality and its relation to network performance than using traditional RTP/UDP. The QoE for the former is still understood to a lesser extent than the latter when it comes to packet loss and other network impairments. Additionally, the overhead introduced when using HTTP/TCP is approximately twice the media bit rate [6].

Other issues that are usually irrelevant for locally hosted applications become important in a cloud context. For example, geographical location can limit the degree of interactivity achieved, as too-remote users may experience unacceptable latency simply due to their distance from the data center where the service runs. A large number of users in various geographical locations can also affect the overall system requirements in terms of scalability and the speed at which it can be achieved. There are a significant number of similar factors that affect all cloud services to some degree, but defining exactly how they impact the end users' QoE remains an open challenge.

Table 1 summarizes the QoE challenges for multimedia cloud applications related to features emerging or changing due to the cloud setting.

QoE for Other Cloud Services

Going beyond multimedia applications, the impact of moving to the cloud becomes more challenging to define. The first challenge is a rather basic one: how can we define what quality of experience means, in the context of any given service? And once we understand what quality means, how can we understand the effect on it when running that particular service on the cloud?

While these questions may sound trivial, or too philosophical for practical consideration, they do pose a significant challenge. The current main driver of cloud services adoption is the expectation of lower costs, for both the service providers and the users. The adoption of services, however, is predicated on said services satisfying the users' expectations in relation to the price paid or, in other words, their utility. If the services do not perform well enough for the users to be satisfied, they will simply not be adopted.

In the case of business applications, service level agreements (SLAs) often need to be in place. SLAs govern how the services should perform, according to the existing contract between the service provider and the customer. In many cases, that performance is related (even if often not explicitly so) to QoE. For example, consider conferencing bridges when a company might have their own private branch exchange (PBX, i.e., a telephone exchange that serves a particular office) or contract a conferencing system from a traditional telecom operator. In those cases, SLAs may exist between the customer and the operator (for either the lines to the PBX or the conferencing service itself), and they can be more or less easily defined. If the same service is considered in a cloud context, however, there are usually more parties involved (e.g., the IaaS provider, the service provider, and a number of network providers between them and the cus-

Going beyond multimedia applications, the impact of moving to the cloud becomes more challenging to define. The first challenge is a rather basic one: how can we define what quality of experience means in the context of any given service?

Feature emerging/changing due to cloud setting	Related QoE challenges from a multimedia perspective
Increase of network distance between end-user and service	Impact of startup delays and waiting times until service is set up on QoE. For media applications, this translates into longer delays and jitter and decreased interactivity in for example VoIP or video-conferencing applications. Cloud-based gaming will also suffer significantly due to higher delays.
Dynamic cloud stack and resource management	Time-varying system conditions require time-dependent QoE models. It is well known, for example, that overall perceptual quality in media applications is subject to a so-called recency effect.
Flexibility to compose new services	Novel influence factors on QoE e.g. viewing angle for personalized live 3D video streaming, or perceived presence for social-IPTV applications.
Distributed cloud	Distributed processing of data may cause novel kinds of artifacts and QoE issues such as temporal inconsistencies. In the case of media-applications sensitive to synchronization issues this may become an issue (e.g. game state in MMORPGs (Massively Multiplayer Online Role-Playing Game) with thousands of concurrent users on several geographically disperse servers).
Service delivery implementation	Technology-driven impairments require new QoE models, e.g. stalling instead of video quality degradation, dealing with latency in cloud gaming.
Multi-party sharing and communication	Interaction between different end-users (e.g. audio- and video-conferencing, collaborative document editing) impose real-time requirements and necessitate new measures that quantify impact of interactivity on QoE
Advanced multi-modal interfaces	Cloud-based processing enables interfaces e.g. based on combined voice and gestural interaction, requiring dedicated multi-model QoE models and metrics as well as mechanisms for distributed processing of sensor data-streams
SLAs and pricing	User expectations as influenced by SLAs and price levels affect user sensitivity and tolerance which need to be reflected by QoE models. For example, for on-demand IPTV services, different guarantees regarding video quality can be offered at different pricing tiers. This concept has no parallel in traditional media delivery channels such as cable TV, where pricing tiers relate only to content.

Table 1. *QoE challenges for multimedia services in a cloud setting.*

tomer). Defining a meaningful SLA then becomes significantly more complex. Defining acceptable performance levels for such a system also becomes more complex. In such cases, having proper definitions and ways to measure QoE could help mitigate this complexity, especially when we consider non-media applications.

Existing Efforts on Cloud QoE

The current research efforts on cloud QoE focus mainly on multimedia applications, in particular on (3D) video streaming, as in the Qualinet community (www.qualinet.eu). The impact of waiting times on user perception, as observed for HTTP streaming like Netflix and YouTube [7], is gaining momentum within the research community due to the increasing popularity of these multimedia clouds. The impact of waiting times is not restricted to multimedia applications only, but can also be applied to interactive services like web browsing [8]. This is caused by fundamental underlying laws for QoE, which is here the well-known Weber-Fechner Law from psychophysics considering waiting times as stimulus. Another generic relationship between QoE and QoS disturbances is quantified by the IQX hypothesis in [9], also applicable to cloud QoE. For non-multimedia applications like remote desktop, there are only singular works available which consider, say, the impact of the network on QoE of Citrix thin clients [10].

The QoE-based classification presented above shows, however, different dimensions beyond multimedia intensity: the level of interactivity, the service complexity, and the usage domain. The usage domain strongly influences the expectations of its users, coming along with SLAs especially for business cloud applications. This may have a strong influence on QoE and has to be reflected accordingly by QoE models; however, there are no works existing so far that relate expectations to QoE. Another open issue is the impact of interactivity of users and its influence on QoE. For example, while audio listening QoE models are well investigated so far, there is only a little literature available on conversational audio QoE models or audio-conferencing QoE models. For complex cloud services like office products, collaborative editing, or operating systems running in the cloud, QoE research is just starting to touch on these challenging areas.

Challenges of QoE Management for Cloud Applications

It is expected that cloud applications and services — including multimedia-intensive applications — will dominate the traffic share worldwide. In this section we first review QoE management in general followed by a detailed discussion of cloud QoE management challenges.

QoE Management in General

QoE management requires three basic steps:
- Understanding and modeling QoE
- Monitoring and estimating QoE
- Adapting and controlling QoE

Understanding and Modeling QoE — Understanding the applications' requirements and the impact of disturbances on the user perceived quality calls for QoE models for given applications, and mappings between measurable parameters and QoE. However, the understanding of QoE still remains a topic of (future) research specifically within cloud applications where new challenges emerge as discussed in the previous section. Hence, generic relationships between measurable parameters and QoE are a fundamental step toward understanding QoE. A typical approach for assessing QoE is calculating mean opinion scores (MOSs) out of subjective tests. That is, the opinions of individual users are aggregated and meant to reflect the opinion of an average user. Due to exponential [9] or logarithmic [11] interdependencies of QoE and QoS, the QoE might be sensitive in certain areas, indicating that MOS cannot solely be used for QoE management. Thus, also user diversity (e.g., reflected by standard deviation of MOS or in terms of distributions) also needs to be taken into account. A generic dependency between user diversity and MOS is proposed in [7].

Monitoring and Estimating QoE — Monitoring includes the retrieval of information such as:
- The network environment (e.g., fixed or wireless)
- The network conditions (e.g., available bandwidth, packet loss)
- Terminal capabilities (e.g., CPU power, resolution, codec)
- SLAs with the network or service operator
- Service- and application-specific information (e.g., content bit rate, genre)

These parameters — if available — are mapped to QoE based on the monitoring done within the network, at the end user, or a combination thereof. While the monitoring within the network can be done by the provider for fast reaction on degrading QoE, it requires mapping functions between network QoS and QoE. When taking into account application-specific parameters additional infrastructure like deep packet inspection (DPI) is required to derive and estimate these parameters within the network. Alternatively, monitoring at the end-user gives the best view on user perceived quality. However, additional challenges arise, e.g., how to feed QoE information back to the provider for adapting and controlling QoE. Hence, trust and integrity issues are critical as users may cheat to get better performance.

Adapting and Controlling QoE — The final step of QoE management is the dynamic adaptation and control thereof to deliver optimal QoE that the user may not get dissatisfied or leave the service. QoE control aims at reacting before the user reacts and uses monitoring information to adjust corresponding impact factors. Open questions in this context are:
- Where to react: at the edge, within the network, or both
- When to react and how often
- How to react and which control knobs to adjust

The final step of QoE management is the dynamic adaptation and control thereof to deliver optimal QoE that the user may not get dissatisfied or leave the service. QoE control aims at reacting before the user reacts and uses monitoring information to adjust corresponding impact factors.

QoE Management for Cloud Applications

For QoE management of cloud applications and services, additional challenges emerge. The basic problem is that a cloud application ecosystem involves more players such as IaaS, PaaS, and SaaS providers in addition to ISPs and telcos. In practice, however, no information is exchanged among the different players to implement QoE management, which covers both QoE monitoring and QoE control. This problem of information asymmetry was already identified in the context of P2P overlay networks. As the structure of the overlay, like in BitTorrent, determines the traffic flows in ISP networks, an ISP can reduce costs due to interdomain traffic and avoid congestion in its own network by influencing the overlay configuration based on information of its own network structure and load. Thus, the ISP has to provide the information to the overlay, which may be utilized accordingly. The key to successful cooperation, however, is incentives for both players, the ISP and the overlay provider. While cost reduction and efficient usage of resources is the major incentive for the ISP, the overlay provider and its users are interested mainly in good QoE. This incentive-based approach to controlling and managing Internet traffic is referred to as economic traffic management (ETM) [12].

Cloud applications pose a similar problem for ISPs since they basically form an overlay between the end user and various data centers. Many cloud applications run distributed across several data centers, and the ISPs have no influence on to which data center a specific user connects. Currently, cloud providers and content delivery networks perform their own traffic optimization despite often being unaware of the actual network conditions and location of end users. This is realized, for example, by redirecting requests to servers "close" to the end user. Nevertheless, the selection of close servers may not be optimal and thus does not necessarily lead to optimal QoE. Additionally, the QoE of cloud applications strongly depends on network conditions and SLAs on the path between the data center and the end user, crossing several different administrative domains. Hence, for successful QoE management, the relevant information has to be made available such that the overlay may be adapted according to the benefit of all players and the users' QoE. An illustration for QoE management of cloud applications is depicted in Fig. 2. Other information possibly exchanged between end user, network provider, and cloud provider lead to improved QoE management (Fig. 3).

In a cloud environment, where there may be several providers involved in delivering a service to a customer, clear and meaningful SLAs can ensure that each party is actually delivering what is expected. In particular, it is interesting to consider the composition of SLAs in terms of the quality of the final service being delivered, which is itself composed of many different parts (computing infrastructure, software platform, network, etc.). This can greatly simplify reasoning about how to optimize the services' QoE by optimizing the performance of the underlying components.

This understanding, in turn, leads to better ways to improve the operational efficiency of the service providers. One way to decrease operating costs, while at the same time improving service levels, is to add as much automation as possible to operating support systems. Automated detection and localization of failures, SLA renegotiation, and automated ticketing are just examples of possible ways in which greater efficiency can be achieved.

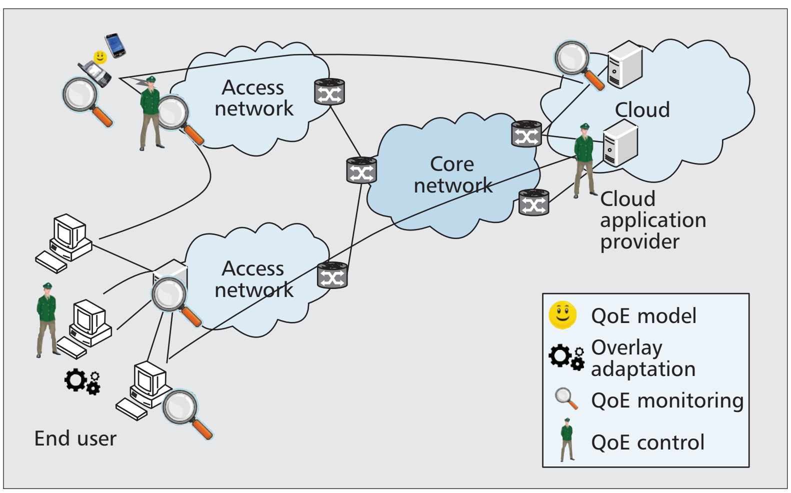

Figure 2. *Prospects for QoE management of cloud applications considering overlay adaptation, QoE monitoring, and QoE control based on a sophisticated cloud QoE model.*

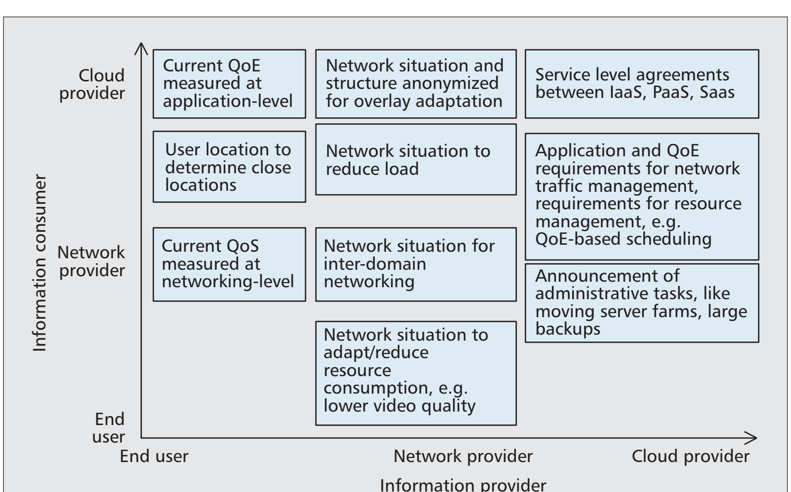

Figure 3. *Interaction between players and exchange of information for improved QoE management.*

QoE Management for Multimedia Cloud Applications

Current approaches to QoE management for multimedia cloud applications focus mainly on QoS provisioning being realized by resource allocation or distributed processing and storage. Hence, QoE models and, in particular, QoS-QoE mappings are required to ensure good user experience followed by the deployment of optimal cloud resource allocation schemes based on, say, service response times and costs of cloud resources [13]. For mobile users, the cloud performs the rendering due to limited computational and power resources on mobile devices (e.g., cloud gaming [5] or cloud-based free viewpoint video/TV [14]). Hence, the resource allocation jointly considers rendering allocation between cloud and client.

Another major challenge is the (distributed) processing of multimedia applications in a cloud. A framework is presented in [15] that proposes a media-edge cloud architecture in which storage, central processing unit (CPU), and graphics processing unit (GPU) clusters are presented at the edge to provide distributed parallel processing and QoS adaptation for various types of devices. However, the focus is still on QoS provisioning, whereas [16] focuses on QoE in distributed environments.

The ALICANTE media ecosystem [17] introduces two novel virtual layers on top of the traditional network layer: a content-aware network (CAN) layer for network packet processing and a Home-Box (HB) layer for the actual content adaptation and delivery. Scalable media coding formats (e.g., H.264/SVC) are deployed for efficient bandwidth-saving delivery of media resources across heterogeneous environments. Cross-layer monitoring probes are distributed within the entire media delivery network (i.e., at content, service, network, end-user premises) providing the required input for advanced QoE management across multiple network domains.

Table 2 summarizes the aforementioned challenges for QoE management of cloud applications and multimedia cloud applications.

Conclusions and Outlook

The shift of computation and data into the cloud has become a key trend in the Internet. With the market approaching perfect competition, the perceived service quality will become an important differentiator between cloud service providers, as the customer is able to choose between different competing providers. For this reason, understanding and managing QoE of end users provides huge opportunities for providers to put themselves at an advantage. It enables cloud providers to observe and react quickly to quality problems, at best before customers perceive them and start churning. From an economic perspective, an optimum QoE has to be achieved while constraining the application to behave as resource-efficiently as possible in order to minimize operational costs. In this article we have shown that economical and efficient use of resources (network, storage, and processing power) while at the same time ensuring suffi-

Topic	QoE management challenge
QoE models for Cloud applications	Not yet mature and currently under research with a new scope of QoE for non-media services.
QoE monitoring and control mechanisms	Depend on the underlying QoE model, for estimating what, where and how to monitor. In an analogous way, QoE control mechanisms have to adequately react to performance issues, in order to maintain the desired QoE levels.
Overlay adaptation as a further step of QoE management for Cloud applications	Depends on the users' location and current situation in the Cloud and in the network with respect to traffic, available resources, etc.
Signaling between network and application to exchange information for QoE management	Requires new interfaces and network entities as discussed in the ALTO group.
Federation between clouds (similar to inter-domain challenges of ISPs)	May be an inhibitor for QoE management, but needs to be realized with open interfaces and common standards.
Development and negotiations of SLAs	May provide the business fundamentals for QoE management.

Table 2. *QoE management challenges for cloud and multimedia cloud applications.*

cient QoE for multiple cloud applications requires comprehensive QoE management solutions. Current approaches, however, are mainly implemented only within the domain of a single stakeholder. Therefore, their effectiveness suffers from an inherent lack of information exchange between the involved constituents, including service infrastructure (IaaS, PaaS, SaaS), network providers (ISPs and telcos), and end users (private and business). To remedy this problem requires flexible cooperation between the entities involved, ultimately enabling every user:

- To access the offered cloud service in any context
- To share content, interact, collaborate, and so on with other users in a dynamic, seamless, and transparent way while maximizing QoE at the same time

To this end, open and standardized interfaces are essential, because only they enable the required content awareness to networks and context awareness to services and applications. This way, the quality and economy of cloud services can be ensured, and thus their widespread and successful adoption.

References

[1] D. Durkee, "Why Cloud Computing Will Never Be Free," Commun. ACM, vol. 53, May 2010; http://doi.acm.org/10.1145/1735223.1735242, pp. 62–69.

[2] ITU-T Rec. P.10/G.100, "Vocabulary and Effects of Transmission Parameters on Customer Opinion of Transmission Quality, Amendment 2," 2006.

[3] M. Creeger, "Cloud Computing: An Overview," Queue, vol. 7, June 2009; http://doi.acm.org/10.1145/ 1551644. 1554608, pp. 2:3–2:4.

[4] L. Qian *et al.*, "Cloud Computing: An Overview," *CloudCom*, *LNCS*, M. G. Jaatun, G. Zhao, and C. Rong, Eds., vol. 5931, Springer, 2009, pp. 626–31.

[5] M. Jarschel *et al.*, "Gaming in the Clouds: Qoe and the Users' Perspective," *Mathematical and Computer Modelling*, no. 0, 2011.

[6] B. Wang *et al.*, "Multimedia Streaming Via TCP: An Analytic Performance Study," *ACM Trans. Multimedia Comp. Commun. Appl.*, vol. 4, May 2008; http://doi.acm.org/10.1145/1352012.1352020, pp. 16:1–16:22.

[7] T. Hoßfeld, R. Schatz, and S. Egger, "SOS: The MOS Is Not Enough!" *QoMEX 2011*, Mechelen, Belgium, Sept. 2011.

[8] S. Egger *et al.*, "Time is Bandwidth? Narrowing the Gap between Subjective Time Perception and Quality of Experience," *IEEE ICC '12*, Ottawa, Canada, June 2012.

[9] M. Fiedler, T. Hoßfeld, and P. Tran-Gia, "A Generic Quantitative Relationship between Quality of Experience and Quality of Service," *IEEE Network* Special Issue on Improving QoE for Network Services, June 2010.

[10] D. Schlosser *et al.*, "Improving the QoE of Citrix Thin Client Users," *IEEE Int'l. Conf. Commun. (ICC 2010)*, Cape Town, South Africa, May 2010.

[11] P. Reichl *et al.*, "The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment," *ICC. IEEE*, 2010, pp. 1–5.

[12] T. Hoßfeld *et al.*, "An Economic Traffic Management Approach to Enable the TripleWin for Users, ISPs, and Overlay Providers," FIA Prague Book. Towards the Future Internet — A European Research Perspective: IOS Press Books Online, ISBN 978-1-60750-007-0, May 2009.

[13] X. Nan, Y. He, and L. Guan, "Optimal Resource Allocation for Multimedia Cloud Based on Queuing Model," *Multimedia and Signal Processing*, 2011.

[14] D. Miao *et al.*, "Resource Allocation for Cloud-Based Free Viewpoint Video Rendering for Mobile Phones," *Proc. 19th Int'l. Conf. Multimedia. ACM*, Nov. 2011, pp. 1237-1240.

[15] W. Zhu *et al.*, "Multimedia Cloud Computing," *IEEE Signal Processing Mag.*, vol. 28, no. 3, May 2011, pp. 59–69.

[16] W. Wu *et al.*, "Quality of Experience in Distributed Interactive Multimedia Environments: Toward A Theoretical Framework," *Proc. 17th ACM Int'l. Conf. Multimedia*, 2009, http://doi.acm.org/10.1145/1631272. 1631338, pp. 481–90.

[17] E. Borcoci, D. Negru, and C. Timmerer, "A Novel Architecture for Multimedia Distribution Based on Content-Aware Networking," *2010 3rd Int'l. Conf. Commun. Theory, Reliability, and Quality of Service*, June 2010, pp. 162–68.

Biographies

TOBIAS HOßFELD (hossfeld@informatik.uni-wuerzburg.de) studied computer science at the University of Würzburg, Germany. He finished his Ph.D. on performance evaluation of future Internet applications and emerging user behavior in 2009. He is now heading the FIA research group "Future Internet Applications & Overlays" at the Chair of Communication Networks in Würzburg. He has been visiting senior researcher at FTW in Vienna with a focus on Quality of Experience research. His main research interests cover social networks, crowd-sourcing, content distribution networks and clouds, as well as modeling Quality of Experience for Internet applications like Skype, YouTube, Web Browsing, and cloud applications in general.

RAIMUND SCHATZ [M] is a senior researcher at the Telecommunications Research Center Vienna (FTW). He holds an M.Sc. in telematics (TU-Graz), a Ph.D. in informatics (TU-Vienna), an M.B.A. (Open University, United Kingdom), and an M.Sc. in international and finance management (Open University). He manages FTW's research projects (GRACE, ACE, ACE 2.0) on quality of experience assessment and monitoring of voice and mobile broadband data services in 3G networks. Being an active member of ACM and IEEE, he has more than 50 publications in the areas of HCI, Quality of Experience and Pervasive Computing.

MARTIN VARELA received his M.Sc. and Ph.D. degrees from the University of Rennes 1, France, in 2002 and 2005 respectively. He also holds a computer engineering degree from the University of the Republic, Uruguay, since 2001. He has worked as a teaching assistant at the University of the Republic and ENST-Bretagne, Rennes, France, and in the Uruguayan Department of Justice as a private contractor. He has been an ERCIM post-doctoral fellow, and has been a visiting researcher at the Swedish Institute of Computer Science and VTT, where he is currently a senior scientist. His main research interests concern the assessment and improvement of the user's QoE in the context of networked multimedia applications and cloud services.

CHRISTIAN TIMMERER is an assistant professor at Alpe-nAdria-Universitt Klagenfurt, Austria. His research interests include the transport and (distributed) adaptation of multimedia content, immersive media technologies, and quality of experience on which he has published more than 80 papers. He participated in several EC-funded projects, notably DANAE, ENTHRONE, P2P-Next, ALICANTE, and SocialSensor. He is an Associate Editor for *IEEE Computing Now*, Area Editor for Elsevier *Signal Processing: Image Communication*, and a Key Member of the Interest Group on Image and Video Coding of IEEE MMTC. He also participated in ISO/MPEG work for several years, notably in the area of MPEG-21, MPEG-M, MPEG-V, and MPEG-DASH.