State of the Art and Research Challenges in the Area of Autonomous Control for a Reliable Internet of Services

Rob van der Mei[1], Hans van den Berg[2(✉)], Ivan Ganchev[3,4,5] ⓘD,
Kurt Tutschku[6], Philipp Leitner[7], Pasi Lassila[8],
Wojciech Burakowski[9], Fidel Liberal[10], Åke Arvidsson[11],
Tobias Hoβfeld[12], Katarzyna Wac[13], Hugh Melvin[14],
Tihana Galinac Grbac[15], Yoram Haddad[16], and Peter Key[17]

[1] Centrum Wiskunde & Informatica (CWI),
Science Park 123, 1098 XG Amsterdam, Netherlands
R.D.van.der.Mei@cwi.nl
[2] TNO, Anna van Buerenplein 1, 2595 DA Den Haag, Netherlands
j.l.vandenberg@tno.nl
[3] University of Limerick, Limerick, Ireland
ivan.ganchev@ul.ie
[4] University of Plovdiv "Paisii Hilendarski", Plovdiv, Bulgaria
[5] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
Sofia, Bulgaria
[6] Blekinge Institute of Technology (BTH),
Campus Grasvik, 37179 Karlskrona, Sweden
kurt.tutschku@bth.se
[7] University of Zurich, Binzmühlestrasse 14, 8050 Zurich, Switzerland
leitner@ifi.uzh.ch
[8] Aalto University, 00076 Espoo, Finland
pasi.lassila@aalto.fi
[9] Institute of Telecommunications, Warsaw University of Technology,
15/19 Nowowiejska Str., 00-665 Warsaw, Poland
wojtek@tele.pw.edu.pl
[10] University of the Basque Country (UPV/EHU), ETSI de Bilbao,
Alameda Urquijo s/n, 48013 Bilbao, Spain
fidel.liberal@ehu.eus
[11] Kristianstad University, Elmetorpsvagen 15, 29188 Kristianstad, Sweden
Ake.Arvidsson@hkr.se
[12] Institute of Computer Science and Business Information Systems (ICB),
Schutzenbahn 70, 45127 Essen, Germany
tobias.hossfeld@uni-due.de
[13] University of Geneva, SES 7, route de Drize, Carouge,
1227 Geneva, Switzerland
Katarzyna.Wac@unige.ch
[14] National University of Ireland - Galway, Galway, Ireland
hugh.melvin@nuigalway.ie
[15] University of Rijeka, Vukovarska 58, 51 000 Rijeka, Croatia
tihana.galinac@riteh.hr

[16] JCT-Lev Academic Center, Guivat Mordehai Neighborhood, PO BOX 16031,
91160 Jerusalem, Israel
haddad@jct.ac.il
[17] Microsoft Research Ltd., Roger Needham Building, 7 J J Thomson Avenue,
Cambridge, UK
peterkey@microsoft.com

Abstract. The explosive growth of the Internet has fundamentally changed the global society. The emergence of concepts like service-oriented architecture (SOA), Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS), Network as a Service (NaaS) and Cloud Computing in general has catalyzed the migration from the information-oriented Internet into an Internet of Services (IoS). This has opened up virtually unbounded possibilities for the creation of new and innovative services that facilitate business processes and improve the quality of life. However, this also calls for new approaches to ensuring quality and reliability of these services. The goal of this book chapter is to first analyze the state-of-the-art in the area of autonomous control for a reliable IoS and then to identify the main research challenges within it. A general background and high-level description of the current state of knowledge is presented. Then, for each of the three subareas, namely the autonomous management and real-time control, methods and tools for monitoring and service prediction, and smart pricing and competition in multi-domain systems, a brief general introduction and background are presented, and a list of key research challenges is formulated.

Keywords: Internet of Services (IoS) · Autonomous control
Autonomous management · Service monitoring · Service prediction
Smart pricing

1   Introduction

Today, we are witnessing a paradigm shift from the traditional information-oriented Internet into an Internet of Services (IoS). This transition opens up virtually unbounded possibilities for creating and deploying new services. Eventually, the Information and Communication Technologies (ICT) landscape will migrate into a global system where new services are essentially large-scale service chains, combining and integrating the functionality of (possibly huge) numbers of other services offered by third parties, including cloud services. At the same time, as our modern society is becoming more and more dependent on ICT, these developments raise the need for effective means to ensure quality and reliability of the services running in such a complex environment.

Motivated by this, the EU COST Action IC1304 "Autonomous Control for a Reliable Internet of Services (ACROSS)" has been established to create a European network of experts, from both academia and industry, aiming at the development of autonomous control methods and algorithms for a reliable and quality-aware IoS.

The goal of this chapter is to identify the main scientific challenges faced during the course of the COST Action ACROSS. To this end, a general background and a

high-level description of the current state of knowledge are first provided. Then, for each of the Action's three working groups (WGs), a brief introduction and background information are provided, followed by a list of key research topics pursued during the Action's lifetime, along with their short description.

2   General Background and Current State of Knowledge

The explosive growth of the Internet has fundamentally changed the global society. The emergence of concepts like service-oriented architecture (SOA), Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS) and Cloud Computing has catalyzed the migration from the information-oriented Internet into an IoS. Together with the Network as a Service (NaaS) concept, enabled through emerging network softwarization techniques (like SDN and NFV), this has opened up virtually unbounded possibilities for the creation of new and innovative services that facilitate business processes and improve the quality of life. As a consequence, modern societies and economies are and will become even more heavily dependent on ICT. Failures and outages of ICT-based services (e.g., financial transactions, Web-shopping, governmental services, generation and distribution of sustainable energy) may cause economic damage and affect people's trust in ICT. Therefore, providing reliable and robust ICT services (resistant against system failures, cyber-attacks, high-load and overload situations, flash crowds, etc.) is crucial for our economy at large. Moreover, in the competitive markets of ICT service offerings, it is of great importance for service providers to be able to realize short time-to-market and to deliver services at sharp price-quality ratios. These observations make the societal and economic importance of reliable Internet services evident.

A fundamental characteristic of the IoS is that services combine and integrate functionalities of other services. This has led to complex service chains with possibly even hundreds of services offered by different third parties, each with their own business incentives. In current practice, service quality of composite services is usually controlled on an ad-hoc basis, while the consequences of failures in service chains are not well understood. The problem is that, although such an approach might work for small service chains, this will become useless for future complex global-scale service chains.

Over the past few years, significant research has been devoted to controlling Quality of Service (QoS) and Quality of Experience (QoE) for IoS. To this end, much progress has been made at the functional layer of QoS-architectures and frameworks, and system development for the IoS. However, relatively little attention has been paid to the development, evaluation and optimization of algorithms for autonomous control that can deal with the growing scale and complexity of the involved service chains. In this context, the main goal of the COST Action ACROSS was to bring the state-of-the-art on autonomous control to the next level by developing quantitative methods and algorithms for autonomous control for a reliable IoS.

In the area of quantitative control methods the main focus has been on 'traditional' controls for QoS provisioning at the network layer and lower layers. In this context, it is important to note that control methods for the IoS also operate at the higher protocol layers and typically involve a multitude of administrative domains. As such, these

control methods – and their effectiveness – are fundamentally different from the traditional control methods, posing fundamentally new challenges. For example, for composite service chains the main challenges are methods for dynamic re-composition, to prevent or mitigate the propagation of failures through the service chains, and methods for overload control at the service level.

Another challenging factor in quality provisioning in the IoS is its highly dynamic nature, imposing a high degree of uncertainty in many respects (e.g., in terms of number and diversity of the service offerings, the system load of services suddenly jumping to temporary overload, demand for cloud resources, etc.). This raises the urgent need for online control methods with self-learning capabilities that quickly adapt to – or even anticipate – changing circumstances [9].

The COST Action ACROSS has brought the state-of-the-art in the area of autonomous quality-based control in the IoS to the next level by developing efficient methods and algorithms that enable network and service providers to fully exploit the enormous possibilities of the IoS. This required conducting a research in the following important sub-areas:

1. Autonomous management and real-time control;
2. Methods and tools for monitoring and service prediction;
3. Smart pricing and competition in multi-domain systems.

These sub-areas were respectively covered by the three ACROSS working groups – WG1, WG2 and WG3. In the following sections, scientific challenges faced in the context of each of these three working groups are elaborated.

3   Autonomous Management and Real-Time Control

On a fundamental level, the working group WG1, associated with this research sub-area, was primarily concerned with the management and control of networks, services, applications, and compositions of services or applications. Of particular interest were management and control techniques that span multiple levels, e.g., the network and service level.

3.1   Introduction and Background

To deliver reliable services in the IoS, service providers need to implement control mechanisms, ranging from simplistic to highly advanced. Typical questions are the following:

- How can one realize the efficient use of control methods by properly setting parameter values and decision thresholds?
- How can one effectively use these mechanisms depending on the specific context of a user (e.g., in terms of user's location, the user's role, operational settings or experienced quality)?
- How do control methods implemented by multiple providers interact?

- How does the interaction between multiple control methods affect their effectiveness?
- What about stability?
- How to resolve conflicts?

Ideally, control mechanisms would be fully distributed and based on (experienced) quality. However, some level of centralized coordination among different autonomous control mechanisms may be needed. In this context, a major challenge is to achieve a proper trade-off between fully distributed control (having higher flexibility and robustness/resilience) and more centralized control (leading to better performance under 'normal' conditions). This will lead to hybrid approaches, aiming to combine 'the best of two worlds'.

3.2   Control Issues in Emerging Softwarized Networks

As part of the current cloud computing trend, the concept of cloud networking [63] has emerged. Cloud networking complements the cloud computing concept by enabling and executing network features and functions in a cloud computing environment. The supplement of computing capabilities to networks outlines elegantly the notion of "softwarization of networks". The added computing capabilities are typically general processing resources, e.g. off-the-shelf servers, which can be used for satisfying computing requirements, i.e. at the application layer (e.g. for the re-coding of videos) or at the network layer (e.g. for the computation of routes). Hence, features and functions in the network-oriented layers are moved away from hardware implementations into software where appropriate, what is lately being termed as network function virtualization (NFV) [24].

The Software-Defined Networking (SDN) paradigm [31] emerged as a solution to the limitations of the monolithic architecture of conventional network devices. By decoupling the system that makes decisions about where traffic is sent (the control plane) from the underlying systems that forward traffic to the selected destination (the data plane), SDN allows network administrators to manage network services through the abstraction of a lower level and more fine-grained functionality. Hence, SDN and the softwarization of networks (NFV) stand for a "new and fine-grained split of network functions and their location of execution". Issues related to the distribution and coordination of software-based network functionality controlling the new simplified hardware (or virtualized) network devices formed a major research issue within ACROSS.

3.3   Scalable QoS-Aware Service Composition Using Hybrid Optimization Methods

Automated or semi-automated QoS-aware service composition is one of the most prevalent research areas in the services research community [25, 56, 85]. In QoS-aware composition, a service composition (or business process, or scientific workflow) is considered as an abstract graph of activities that need to be executed. Concrete services can be used to implement specific activities in the graph. Typically, it is assumed that

there are multiple functionally identical services with differing QoS available to implement each activity in the abstract composition. The instantiation problem is then to find the combination of services to use for each activity so that the overall QoS (based on one or more QoS metrics) is optimal, for instance, to minimize the QoS metric "response time" given a specific budget. The instantiation problem can be reduced to a minimization problem, and it is known to be NP-complete.

Traditionally, QoS-aware service composition has been done using deterministic methods (e.g., simplex) for small service compositions, or a wide array of heuristics for large-scale problem instances (e.g., genetic algorithms, simulated annealing, and various custom implementations). However, the advent of cloud services and SDNs, service brokers, as well as the generally increasing size of service compositions require new hybrid methods, which combine locally optimal solutions on various levels (e.g., the network, application, or service broker level). It is yet unclear how such optimizations on various levels, conducted by various separate entities, can be optimally performed and coordinated, and how stability of such systems can be ensured. However, one promising approach is the utilization of nature-inspired composition techniques, for instance, the chemical programming metaphor [25, 58].

3.4   Efficient Use of Cloud Federation and Cloud Bursting Concepts

One of the challenges of current cloud computing systems is the efficient use of multiple cloud services or cloud providers. On the network level, this includes the idea of virtual network infrastructures (VNIs), c.f. [44]. The VNI concept assumes exploitation of network resources offered by different network providers and their composition into a common, coherent communication infrastructure supporting distributed cloud federation [17]. Controlling, managing, and monitoring network resources would allow cloud federations to implement various new features that could: (1) optimize traffic between sites, services, and users; (2) provide isolation for the whole clouds or even for particular users, e.g. who require deployment of their own protocols over the network layer; (3) simplify the process of extending and integrating cloud providers and network providers into a federation with reduced efforts and costs.

On the service and application level, the idea of cloud bursting has been proposed as a way to efficiently use multiple cloud services [32, 57]. In cloud bursting, applications or services are typically running in a private cloud setup, until an external event (e.g., a significant load spike that cannot be covered by internal resources) forces the application to "burst" and move either the entire application or parts of it to a public cloud service. While this model has clear commercial advantages, its concrete realization is still difficult, as the cloud bursting requires intelligent control and management mechanisms for predicting the load, for deciding which applications or services to burst, and for technically implementing a seamless migration. Additionally, the increased network latency is often a current practical problem in cloud bursting scenarios.

3.5   Energy-Aware Network and Service Control

Traditionally, the optimization of ICT service provision made use of network performance related characteristics or key performance indicators (KPI) as basic inputs for

control and actuation loops. Those initially simple and technical-only parameters evolved later to more complex QoE related aspects, leading to multi-variate optimization problems. New control and actuation loops then involved several parameters to be handled in a joint manner due to the different trade-offs and interdependencies among input and output indicators. This is usually done by composing the effects through a simplified utility function. Therefore, resulting approaches have put the focus particularly in the reward (in terms of users' satisfaction) to be achieved by efficiently using the available network resources (c.f. [75]).

Meanwhile, the cost of doing so was most of the times faced as a constraint of the mathematical problem and considered again technical resources only. However, in the wake of "green ICT" and, more generally speaking, the requirement of economically sustainable and profitable service provision entail new research challenges where the cost of service provisioning must also consider energy consumption and price (c.f. [74]). The resulting energy- and price-aware control loops demand intensive research, as the underlying multi-objective optimization problem as well as the complexity of utility functions (c.f. [60]) and the mechanisms for articulation of preferences exceed current common practices.

Such constraints affect not only the network but also the whole ICT service provision chain. For example, server farms are vital components in cloud computing and advanced multi-server queueing models that include features essential for characterizing scheduling performance as well as energy efficiency need to be developed. Recent results in this area include [29, 40, 41] and analyze fundamental structural properties of policies that optimize the performance-energy trade-off. On the other hand, several works exist [20, 67] that employ energy-driven Markov Decision Process (MDP) solutions. In addition, the use of energy-aware multi-path TCP in heterogeneous networks ([15, 21]) has become challenging.

3.6   Developments in Transport Control Protocols

Transport protocols, particularly TCP and related protocols, are subject to continuous evolution for at least two reasons besides the omnipresent, general desire to improve. The first reason is a need to keep up with the development of internet infrastructure with, *e.g.,* reduced memory costs, widespread fibre deployment and high speed cellular technologies which enable larger buffers, higher bit rates and/or more variable channels. The second reason is the increasing competition between providers of internet based services which drives various efforts to keep ahead of the competition in terms of user experience. The results are new versions of the TCP congestion control algorithm as well as new protocols to replace TCP.

The work on new TCP congestion control algorithms includes work on adapting to the changing characteristics of the internet such as the higher and more variable bandwidths offered by cellular accesses [1, 11, 34, 50, 52, 80, 84], possibly using cross layer approaches [6, 10, 59, 61], but also simple tuning of existing TCP such as increasing the size of the initial window [13, 16, 66, 82].

The efforts to replace TCP include QUIC (Quick UDP Internet Connection, a protocol from Google) and SPUD (Session Protocol for User Datagrams, an IETF initiative) and its successor PLUS (Path Layer UDP Substrate, also an IETF initiative), c.f. [76]. The QUIC protocol aims at reducing latencies by combining connection establishment (three way handshake in TCP) with encryption key exchange (presently a second transaction); it also includes the possibility of completely eliminating key exchange if cached keys are available and can be reused. Another key feature is the built-in support for HTML/2 such that multiple objects can be multiplexed over the same stream [18, 72]. The purpose of SPUD/PLUS is to offer and end-to-end transport protocol based on UDP with support for direct communication with middleboxes (e.g., firewalls). The rationale for this is the difficulties with developing TCP that follow from the fact that present middleboxes rely on implicit interpretations of TCP, and/or lack of encryption to perform different forms of functionality some of which even may be unwanted. Examples of such implicit interpretations include TCP packets with SYN and ACK flags being interpreted by gateways as confirmations of NAT (network address translation) settings and by firewalls as confirmations of user acceptance [23, 49]. Examples of possibly unwanted functionality include traffic management devices aborting flows by manipulating the RST flag in TCP packets [22].

New versions of TCP or new DIY (do-it-yourself) protocols open a world of threats and opportunities. The threats range from unfair competition [18, 72] to the risk of congestion collapse as content providers develop more and more aggressive protocols and deploy faster and faster accesses in an attempt to improve their service [13, 66, 82]. But it also includes the inability to cache popular objects near users or prioritize between flows on congested access links as a result of the tendency to paint all traffic "grey", i.e. to encrypt even trivial things like public information (cf. Section 4.6). As for opportunities, TCP clearly has some performance problems and is a part of the ossification of the Internet. A (set of) new protocol(s) could circumvent the issues related to TCP and be adapted to present networks and content, and therefore provide potentially better performance.

The goal of the work on transport protocols in this context is, primary, to evaluate existing transport protocols and, secondary, to present new congestion control algorithms and/or new transport protocols that work better than present TCP, and at the same time compete with well behaved, legacy TCP in a fair way.

4 Methods and Tools for Monitoring and Service Prediction

Methods and tools for monitoring and service prediction was the main topic of WG2, mostly considered in the context of a larger system that needs to be (autonomously) controlled.

4.1 Introduction and Background

A crucial element for autonomous control in the IoS is monitoring and service prediction. For autonomous real-time (user-perceived) QoS and QoE in large, dynamic, complex multi-domain environments like the IoS, there is a great need for scalable,

non-intrusive monitoring and measurement of service demands, service performance, and resource usage. Additional constraints regarding, for instance, privacy and integrity further complicate the challenges for monitoring and measurement. In addition, proactive service adaptation capabilities are rapidly becoming increasingly important for service-oriented systems like IoS. In this context, there is a need for online quality prediction methods in combination with self-adaptation capabilities (e.g., service re-composition). Service performance monitoring capabilities are also important for the assessment of Service Level Agreement (SLA) conformance, and moreover, to provide accurate billing information. In general, the metrics to monitor rely on the point of view adopted. For instance, cloud providers need metrics to monitor SLA conformance and manage the cloud whereas composite service provider have to monitor multiple SLAs which is also different than what is required to be monitored for customers and service consumers.

4.2   How to Define 'QoS' and 'QoE', and What to Measure?

A common definition of QoE is provided in [55]: "*QoE is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state.*" In contrast, the ITU-T Rec. P.10 defines QoE as "*the overall acceptability of an application or service, as perceived subjectively by the end user*". The definition in [55] advances the ITU-T definition by going beyond merely binary acceptability and by emphasizing the importance of both, pragmatic (utility) and hedonic (enjoyment) aspects of quality judgment formation. The difference to the definition of QoS by the ITU-T Rec. E.800 is significant: *"[the] totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service*". Factors important for QoE like context of usage and user characteristics are not comprehensibly addressed by QoS.

As a common denominator, four different categories of QoE influence factors [37, 55] are distinguished, which are the influence factors on the context, user, system, and content level (Fig. 1). The context level considers aspects like the environment where the user is consuming the service, the social and cultural background, or the purpose of using the service like time killing or information retrieval. The user level includes psychological factors like expectations of the user, memory and recency effects, or the usage history of the application. The technical influence factors are abstracted on the system level. They cover influences of the transmission network, the devices and screens, but also of the implementation of the application itself like video buffering strategies. The content level addresses, for instance on the example of video delivery, the video codec, format, resolution, but also the duration, contents, and type of the video and its motion patterns.
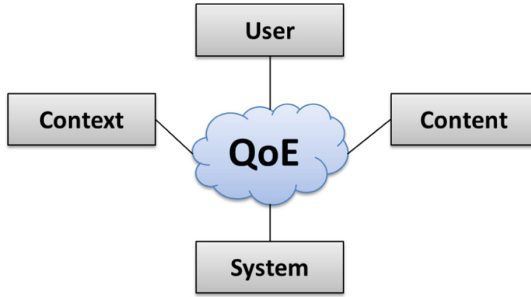
Fig. 1. Different categories of QoE influence factors

4.3   QoE and QoS Monitoring for Cloud Services

The specific challenges of QoE management for cloud services are discussed in detail in [38]. Cloud technologies are used for the provision of a whole spectrum of new and also traditional services. As users' experiences are typically application- and service-dependent, the generality of the services can be considered a big challenge in QoE monitoring of cloud services. Nevertheless, generic methods would be needed as tailoring of models for each and every application is not feasible in practice. Another challenge is brought up by multitude of service access methods. Nowadays, people use variety of different devices and applications to access the services from within many kinds of contexts (e.g. different social situations, physical locations, etc.).

Traditional services that have been moved to clouds can continue using the proven existing QoE metrics. However, new QoS metrics related to the new kind of resources and their management (e.g. virtualization techniques, distributed processing and storage) and how they contribute to QoE require gaining new understanding. On the other hand, the new kind of services enabled by the cloud technologies (e.g. storage and collaboration) call for research regarding not only QoS-to-QoE mapping, but also the fundamentals on how users perceive these services. In addition to this, the much discussed security, privacy, and cost need to be considered inside the QoE topic.

4.4   QoE and Context-Aware Monitoring

Today's consumer Internet traffic is transmitted on a best effort basis without taking into account any quality requirements. QoE management aims at satisfying the demands of applications and users in the network by efficiently utilizing existing resources. Therefore, QoE management requires an information exchange between the application and the network, and proper monitoring approaches. There are three basic research steps in the QoE management: (1) QoE modeling; (2) QoE monitoring; and (3) QoE optimizing.

As a result of the QoE modeling process, QoE-relevant parameters are identified which have to be monitored accordingly. In general, monitoring includes the collection of information such as: (1) the network environment (e.g., fixed or wireless); (2) the network conditions (e.g., available bandwidth, packet loss, etc.); (3) terminal

capabilities (e.g., CPU, memory, display resolution); (4) service- and application-specific information (e.g., video bit rate, encoding, content genre) [26, 69]. But also monitoring at the application layer may be important. For instance, QoE monitoring for YouTube requires monitoring or estimating the video buffer status in order to recognize or predict when stalling occurs.

The QoE monitoring can either be performed: (1) at the end user or terminal level; (2) within the network; or (3) by a combination thereof. While the monitoring within the network can be done by the provider itself for a fast reaction to QoE degradation, it requires mapping functions between network QoS and QoE. When taking into account application-specific parameters additional infrastructure like deep packet inspection (DPI) may be required to derive and estimate these parameters within the network. A better view on user perceived quality is achieved by monitoring at the end user level. However, additional challenges arise, e.g., how to feed QoE information back to the provider for adapting and controlling QoE. In addition, trust and integrity issues are critical as users may cheat to get better performance [68].

Going beyond QoE management, additional information may be exploited to optimize the services on a system level, e.g. allocation and utilization of system resources, resilience of services, but also the user perceived quality. While QoE management mainly targets the optimization of current service delivery and currently running applications, the exploitation of context information by network operators may lead to a more sophisticated traffic management, a reduction of the traffic load on inter-domain links, and a reduction of the operating costs for the Internet service providers (ISPs).

Context monitoring aims at getting information about the current system situation from a holistic point of view. Such information is helpful for control decisions. For example, the popularity of video requests may be monitored, events may be foreseen (like soccer matches) which allow to better control service and allocate resources. This information may stem from different sources like social networks (useful for figuring out the popularity of videos and deciding about caching/bandwidth demands) but also can be monitored on the fly. Thus, context monitoring includes aspects beyond QoE monitoring (Fig. 2) [39]. Context monitoring increases QoS and QoE (due to management of individual flows/users). But it may also improve the resilience of services (due to broad information about the network "status") [64].
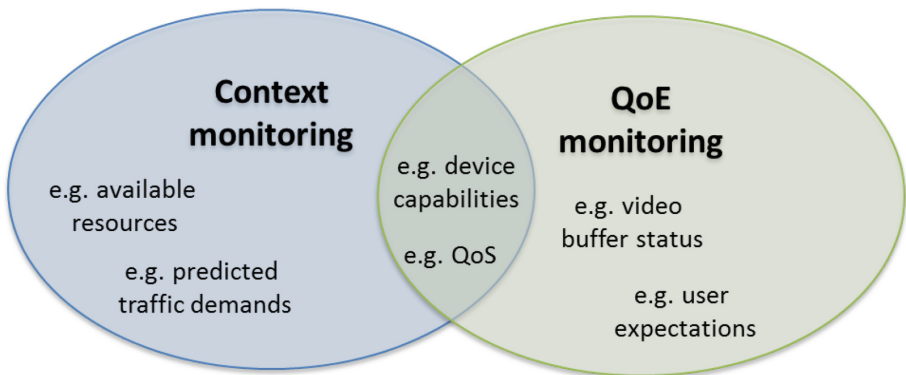

Fig. 2. Relation between QoE monitoring and context monitoring

Context monitoring requires models, metrics, and approaches which capture the conditions/state of the system (network infrastructure, up to and including the service layer), but also application/service demands and the capabilities of the end-user device. The challenges here are the following: (1) identification of relevant context information required for QoE but also reliable services; (2) quantification of QoE, based on relevant QoS and context information; and (3) monitoring architecture and concept.

4.5 Inclusion of Human Factors

Inevitably, Internet applications and services on a growing scale assist us in our daily life situations, fulfilling our needs for leisure, entertainment, communication or information. However, on one hand, user acceptance of an existing Internet service/application depends on the variety of human factors influencing its perception, and, on the other hand, there are many human factors and needs, which could be supported by the Internet services and computing at large, yet unknown to date. However, despite the importance of understanding of the human factors in computing, a sound methodology for evaluation of these factors and delineation of new ones, as well as reliable methods to design new Internet services with these factors in mind, do not exist.

This challenge goes beyond the QoE/QoS challenge presented in the previous subsection relating to the user experience with respect to an existing and used system. The challenge presented here relates to identification of the unmet (implicit) needs of the user enabling future provision of novel and useful services. These human factors may relate to some specific phenomena ranging from, for example, the most preferred interaction style with a service (e.g., auditory, kinesthetic, visual) in a given context, via the user's specific health and care needs (e.g., wellness or anti-ageing), to the user's specific factors like cognitive load, physical flexibility, or momentary perception of safety, or intimacy in a specific context [33, 42, 54]

In this challenge, one aims to provide a set of rigorous interdisciplinary, i.e., mixed-methods based methodological steps to be taken aiming to quantify human factors in computing within the user's natural environments and different contexts of service usage [78]. The methodology incorporates qualitative and quantitative methods and involves real users in their real life environments through:

- Gathering the cumulative users' opinion via open-ended interviews and surveys. Thus, specifically focusing on understanding the users' expectations towards a researched phenomenon and current experience of this phenomenon, mostly to establish the users' baseline experience on the experiment variables and context, but also to gather general demographics about the experiment participants.
- Gathering the momentary users' opinion upon some specific factors like health behaviors, moods, feelings, social interactions, or environmental and contextual conditions via the Experience Sampling Method (ESM). Special momentary surveys executed multiple times per day 'in situ', i.e., in the natural users' environments [79].
- Gathering the episodic users' opinion upon some specific factors (as above) along semi-structured interviews based on the diary, for example by the Day Reconstruction Method.

- Gathering the data upon the users' daily life contexts and smartphone usage via continuous, automatic, unobtrusive data collection on the users' device through the measurements-based 'Logger' service.

Secondly, in this challenge, one wish to provide guidelines on analyzing the relation of these factors with the design features of the computing system itself. Thirdly, one would like to provide guidelines for Internet services/applications leveraging the human factors in their design process, assuring user's experience (QoE) and thus maximizing the user acceptance for these services.

4.6   Aggregated and Encrypted Data ('Grey Traffic')

Monitoring generally assumes that it is possible to extract from the data a set of parameters (e.g. fields within a packet) that enables to know what data (or services) is travelling (resp. provided). However, there is recent tendency to paint all traffic "grey", i.e. to encrypt even trivial things like public information. Even though this may appear to protect user privacy, in fact such obfuscation complicates or prevents monitoring, caching, and prioritization which could have been used to reduce costs and optimize user experience. Actually, it is not only the content that is being encrypted but also the protocol itself (i.e. only the UDP header or similar is left open). This means that, contrary to present TCP, one cannot even monitor a flow in terms of data and acknowledgments to, e.g. detect malfunctioning flows (e.g., subject to extreme losses) or perform local retransmission (e.g. from a proxy). Regarding content identification, the solution needs not necessarily be unprotected content (there are reasons related to content ownership, etc.) but one can imagine tags of different kinds. Then there is a challenge to find incentives that encourage correct labelling [12], such that traffic can be monitored and identified to the extent necessary to optimize networks (long term) and QoE (short term).

4.7   Timing Accuracy for Network and Service Control

A key objective of ACROSS was to ensure that the ICT infrastructure that supports future Internet is designed such that the quality and reliability of the services running in such a complex environment can be guaranteed. This is a huge challenge with many facets, particularly as the Internet evolves in scale and complexity. One key building block that is required at the heart of this evolving infrastructure is precise and verifiable timing. Requirements such as 'real-time control', 'quality monitoring', 'QoS and QoE monitoring', 'SDN' cannot easily or effectively be met without a common sense of precise time distributed across the full infrastructure. 'You cannot control what you do not understand' is a phrase that applies here – and you cannot understand a dynamic and real-time system without having precise and verifiable timing data on its performance. As a first step, such timing services will firstly ensure that application- and network performance can be precisely monitored, but secondly, and more importantly, will facilitate the design of better systems and infrastructures to meet the future needs. Unfortunately, current ICT systems do not readily support this paradigm [83]. Applications, computers and communications systems have been developed with modules

and layers that optimize data processing but degrade accurate timing. State-of-the-art systems now use timing but only as a performance metric. To enable the predicted massive growth, accurate timing needs cross-disciplinary research to be integrated into existing and future systems. In addition, timing accuracy and security of timing services issues represent another critical need. In many cases, having assurance that the time is correct is a more difficult problem than accuracy. A number of recent initiatives are focusing on these challenges, c.f. [19, 43, 71, 73, 81].

4.8   Prediction of Performance, Quality and Reliability for Composite Services

Service orientation as a paradigm which switches software production to software use is gaining more popularity; the number of services is growing with higher potential of service reuse and integration into *composite services*. There may be number of possibilities to create specific composite services and they may differ in structure and selection of services that form this composition. The composite services are also characterized by functional and non-functional attributes. Here the focus is on prediction models for behavior of composite services in respect to performance, quality and reliability. There are many approaches to build Quality of Service (QoS) aware service compositions [47] and most of them are based on heuristic and meta-heuristic approaches. However, there is a lack of mathematical models that provide better understanding of underlying causes that generate particular QoS behavior.

Regarding QoS and reliability prediction of composite services, it is well known that size-, fault-and failure distribution over software components in large scale complex software systems follow power law distributions [30, 36]. The knowledge of underlying generative models for these distributions enables developers to identify critical parts of such systems at early stages of development and act accordingly to produce higher quality and more reliable software at lower cost. Similar behavior is expected from large-scale service compositions. The challenge is to extend the theory of distribution of size, faults and failures to other attributes of services (e.g. above mentioned non- functional attributes) in large-scale service compositions. Identification of such distributions, that may contain generative properties, would enable to predict the behavior of composite services.

4.9   Monitoring with SDN

SDN is a new and promising networking paradigm [53, 62]. It consists in decoupling control plane from forwarding plane and offers a whole set of opportunities to monitor the network performance. In SDN, each node (router, switch, …) updates a controller about almost any information regarding the traffic traveling at any time in the network. A set of patterns can be defined by the controller for the node to apply and count the number of packets matching this specific pattern. Basic monitoring applies on the well-known header fields at any communication layer. However, some NFV can be introduced at some nodes to perform fine monitoring on data (e.g. DPI to get specific info from data) and therefore to enable the controller to have full knowledge on what happens in the network.

Of course these great opportunities provided by SDN are accompanied by a list of (measurement and monitoring) challenges currently researched over the world [14, 35, 46]. For instance, how many controllers should be deployed? Too many controllers would bring us back to the older architecture but on the other hand too few controllers that centralize a large area would induce delay in getting the information and would require very expensive computation power to deal with huge amount of data. In this latter case, this would also generate bottleneck near the area of the controller(s).

5   Smart Pricing and Competition in Multi-domain Systems

WG3 dealt with pricing and competition in the IoS, in particular in relation to service quality and reliability.

5.1   Introduction and Background

Service providers in the IoS could implement their own pricing mechanism, which may involve simple static pricing to advanced dynamic policies where prices may e.g. vary (even at small time scale) according to the actual demand [4]. The involvement of third-party and cloud services in making up a composite service in these dynamic and competitive environments (with all involved parties striving for maximization of their own profit) raises challenging questions that are new, even though one can learn from the past. For example, in the traditional Internet, volume-based charging schemes tend to be replaced by flat-fee charging schemes. In this context, typical questions are: (1) what are the implications of implementing different pricing mechanisms in a multi-domain setting? (2) how do quality levels and pricing mechanisms relate? (3) how can one develop smart pricing mechanisms that provide proper incentives for the involved parties (regarding brokering, SLA negotiation strategies, federation, etc.) that lead to a stable ecosystem? (4) what governing rules are needed to achieve this?

5.2   Modeling QoS/QoE-Aware Pricing Issues

A key challenge is to understand what are the correct digital "goods" (e.g. in the cloud, in a distributed setting, beyond just physical resources), and at what level of granularity to consider pricing and competition issues [2, 7]. An overview of some of the pricing issues for the cloud is given in [51]. Initial cloud services were primarily resource based, with different types of resources (such as compute power, storage, bandwidth), different types of service (service and batch) and different service types (IaaS, SaaS etc.). Simple fixed pricing schemes are typically used by the providers, with the large cloud providers forming an oligopoly and competing on price. But even in this setting, each of the individual component resources and services have their own QoS measures and SLAs, which makes specifying the QoS and QoE of an actual service used by a customer difficult. The landscape is also changing: different types of cloud service providers are emerging, as are different types of services (such as data services, data

analytics, automated Machine Learning), which brings additional complexity. Hence research is needed on the following subtopics:

- *The digital goods and services for an IoS*. The challenge is to identify the fundamental building blocks: for example, are they just physical or virtual resources, as with current IaaS/PaaS, or do they include abstract computing elements and capabilities? Can they include software abstractions that would enable flexible SaaS descriptions rather than the current, limited, application specific SaaS offering? How can data be included as a good? Can higher layer services, such as automated analytics or machine learning be specified as capabilities? A fundamental question for pricing is whether goods and services are multidimensional or can be thought of a primarily unidimensional (see [51]).
- *A QoS and QoE framework for describing services*. The current state of the art in IaaS is for providers to specify individual resources or bundles, each with some QoS measure or SLA, that often is just based on availability or mean throughput. The customer has to decide what to purchase and assemble the different resources and, somehow, translate into the solution into a QoS or QoE for their usage scenario. At the other extreme, specific solutions are offered by SaaS for limited applications (e.g. SAP). As the service and solutions that customers need or want to offer to their own customers become ever richer, a framework is needed that allows realistic services to be described in terms of their own QoS and QoE.
- *Component specification that allows services to be built up from components*. The challenges here are closely tied to those for QoS and QoE. The current bottom-up purchase and construction of services from individual components makes life easy for providers but difficult for customers and solution providers, who would typically want a top-down specification. For example, an end-customer may see their data as the primary resource, building services and analytics based on it, and hence want performance and QoS measures related to these. There is a need to be able to build services from different components and different providers; the challenge is how to achieve this.
- *Brokering, transfer charging and "exchanges" to allow for third parties, and for multi-provider services*. Pricing models in use now are basic: they typically involve pay-as-you-go pricing, with discounts for bundling, and with a rudimentary reservation offering. Amazon offers a Spot market for IAAS, although the pricing doesn't appear to reflect a true auction mechanism [5]. There is a need for more flexible pricing models to enable users with flexible workloads to balance price against performance, and to reflect elastic demand. Research is needed to see how transfer charging may encourage multi-provider services, and whether compute and data resources can be treated as digital commodities and traded in exchanges.

5.3   Context-Dependent Pricing, Charging and Billing of Composite Services

Pricing, charging and billing of composite services, provided to the end user by different service providers, require the construction and elaboration of new mechanisms and techniques in order to provide the best service to the user [3, 48, 77], depending on

their current context[1], and to enable viable business models for the service providers [8, 28, 65]. Solving this problem requires advances in mechanism design: current economic theory lacks the sophistication to handle the potentially rich variety of service descriptions and specifications that could be delivered in the IoT and Next Generation Internet (NGI).

Charging and billing (C&B) requires mechanisms to allow secure transactions, trusted third party (TTP) C&B [45], cross-provider payments, micropayments and to allow for new payment paradigms, such as peer-to-peer currencies. The TTP feature of the C&B entity, perhaps, will also facilitate the initial establishment of trust and subsequent interaction, e.g. to ensure interoperability, between different service providers as regards the services (service components) provided by each of them.

The pricing and C&B need to be aligned with service definition and implementation. Hence the autonomous control aspects (ACROSS WG1) need to be inextricably linked to pricing, and what can be measured (ACROSS WG2). This challenge relates also to the services' intelligent demand shaping (IDS) and services' measurement, analytics, and profiling (MAP).

As a specific example, service delivery and SLAs are linked to the dynamic monitoring of the quality of each component of the composite service, with an ability to dynamically replace/substitute the component(s) that is/are currently underperforming with another one(s), which is/are identified as working better in the current context. The replacement of service components must be performed transparently to the user – perhaps with the user only noticing improvements in the overall service quality.

5.4 QoS and Price-Aware Selection of Cloud Service Providers

The upraise of IaaS clouds has led to an interesting dilemma for software engineers. Fundamentally, the basic service offered by different providers (e.g., Amazon EC2, or, more recently, Google Compute Engine) is entirely interchangeable. However, non-functional aspects (e.g., pricing models, expected performance of acquired resources, stability and predictability of performance) vary considerably, not only between providers, but even among different data centers of the same provider. This is made worse by the fact that, currently, IaaS providers are notoriously vague when specifying details of their service (e.g., "has two virtual CPUs and medium networking performance"). As a consequence, cloud users are currently not able to make an informed decision about which cloud to adopt, and which concrete configuration (e.g., instance type) to use for which application. Hence, cloud users often base their most fundamental operations decisions on hearsay, marketing slogans, and anecdotal evidence rather than sound data. Multiple research teams worldwide have proposed tools to allow developers to benchmark cloud services in a more rigid way prior to

---

[1] The *context*, from which the price is computed, has three main components – *user context* (i.e. the user's location, preferences and profile(s), the user mobile device(s), etc.), *network context* (i.e. the congestion level, the current data usage pattern, the current QoS/QoE index, the cost of using a network, etc.), and *service context* (i.e. the category, type, scope, and attributes of the service being requested, the request time, the application initiating the request, the current QoS/QoE index of the service component, etc.).

deployment (e.g., CloudCrawler, CloudBench, or Cloud Workbench, [27, 70]). However, so far, fundamental insights are missing as regards which kind of IaaS provider and configuration is suitable for which kind of application and workload.

6   Conclusion

As can be seen from this chapter, there is a high variety of research challenges in the area of autonomous control for a reliable Internet of Services (IoS), which of course cannot be covered by a single book. The following chapters deal with a subset of these, mainly related to service monitoring, control, management, and prediction, leaving the rest of challenges for another book.

References

1. Abdelsalam, A., Luglio, M., Roseti, C., Zampognaro, F.: TCP wave: a new reliable transport approach for future internet. Comput. Netw. **112**, 122–143 (2017)
2. Abhiskeh, V., Kash, I.A., Key, P.: Fixed and market pricing for cloud services. In: Proceedings of 7th Workshop on Economics of Networks, Systems and Computation (NetEcon 2012), Orlando, USA (2012)
3. Abundo, M., Valerio, V.D., Cardellini, V., Lo Presti, F.: Bidding strategies in QoS-aware cloud systems based on N-armed bandit problems. In: Proceedings of the 3rd IEEE Symposium on Network Cloud Computing and Applications (NCCA 2014), pp. 38–45 (2014)
4. Alzhouri, F., Agarwal, A., Liu, Y., Bataineh, A.S.: Dynamic pricing for maximizing cloud revenue: a column generation approach. In: Proceedings of the 18th International Conference on Distributed Computing and Networking (ICDCN 2017) (2017). Article No. 22
5. Ben-Yehuda, O.A., et al.: Deconstructing amazon EC2 spot instance pricing. ACM Trans. Econom. Comput. **1**(3), 16:1–16:20 (2013)
6. Beshay, J.D., Taghavi Nasrabadi, A., Prakash, R., Francini, A.: Link-coupled TCP for 5G networks. In: IEEE/ACM 25th International Symposium on Quality of Service (IWQoS), pp. 1–6 (2017)
7. Bhattacharjee, S., Gopal, R.D., Marsden, J.R., Sankaranarayanan, R.: Digital goods and markets: emerging issues and challenges. ACM Trans. Manag. Inf. Syst. **2**(2), 1–23 (2011)
8. Blocq., G., Bachrach, Y., Key, P.: The shared assignment game and applications to pricing in cloud computing. In: AAMAS 2014, Paris, France (2014)
9. Bosman, J., Van den Berg, H., Van der Mei, R.: Real-time QoS control for service orchestration. In: Proceedings of the 27th International Teletraffic Congress (ITC27), pp. 152–158 (2015)
10. Cabrera Molero, I., Möller, N., Petersson, J., Skog, R., Arvidsson, Å., Flärdh, O., Johansson, K.H.: Cross-layer adaptation for TCP-based applications in WCDMA systems. In: IST Mobile and Wireless Communications Summit, Dresden, Germany (2005)
11. Cardwell, N., Cheng, Y., Gunn, C.S., Yeganeh, S.H., Jacobson, V.: BBR: congestion-based congestion control. ACM Queue **14**(5), 20–53 (2016)

12. Casaretto, J.: The Internet strikes back: global encrypted SSL traffic booms (2014). http://siliconangle.com/blog/2014/05/20/the-internet-strikes-back-global-encrypted-ssl-traffic-booms/
13. CDNPlanet Initcwnd settings of major CDN providers (2017). https://www.cdnplanet.com/blog/initcwnd-settings-major-cdn-providers. Accessed 12 Oct 2017
14. Chaudet, C., Haddad, Y.: Wireless software defined networks: challenges and opportunities. In: IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems (COMCAS 2013), pp. 1–5 (2013)
15. Chen, S., Yuan, Z., Muntean, G.M.: A traffic burstiness-based offload scheme for energy efficiency deliveries in heterogeneous wireless networks. In: IEEE Globecom Workshops, Atlanta, USA (2013)
16. Chu, J., Dukkipati, N., Cheng, Y., Mathis, M.: Increasing TCP's Initial Window. IETF RFC 6928, April 2013
17. CONTRAIL project (Seamless cloud federation through open virtualization infrastructure). In: EU 7th Framework Programme. http://contrail-project.eu/
18. Cook, S., Mathieu, B., Truong, P., Hamchaoui, I.: QUIC: better for what and for whom? In: IEEE International Conference on Communications (ICC), pp. 1–6 (2017)
19. Cyber-Physical Systems Public Working Group (CPS PWG). http://www.cpspwg.org/
20. Dechen, D., Shami, A.: Energy efficient QoS constrained scheduler for SC-FDMA uplink. Pers. Commun. **8**, 81–90 (2013)
21. Ding, T., et al.: Smartphone energy consumption models for multimedia services using multipath TCP. In: Proceedings of the 11th IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, USA (2014)
22. Eckersley, P., Von Lohmann, F., Schoen, S.: Packet forgery by ISPs: a report on the comcast affair. In: Electronic Frontier Foundation (2007). https://www.eff.org/sv/wp/packet-forgery-isps-report-comcast-affair. Accessed 12 Oct 2017
23. Edeline, K., Kühlewind, M., Trammell, B., Aben, E., Donnet, B.: Using UDP for internet transport evolution. ETH TIK Technical Report 366 (2016)
24. ETSI ISG on NFV (2014) Network functions virtualisation – white paper #3. http://portal.etsi.org/NFV/NFV_White_Paper3.pdf
25. Fernández, H., Tedeschi, C.: A chemistry-inspired workflow management system for decentralizing workflow execution. IEEE Trans. Serv. Comput. **9**(2), 213–226 (2016)
26. Fiedler, M., Hoßfeld, T., Tran-Gia, P.: A generic quantitative relationship between quality of experience and quality of service. IEEE Netw. **24**(2), 36–41 (2010)
27. Frey, S., Fittkau, F., Hasselbring, W.: Search-based genetic optimization for deployment and reconfiguration of software in the cloud. In: 2013 International Conference on Software Engineering (ICSE 2013), pp. 512–521. IEEE Press, Piscataway (2013)
28. Gao, X.A., Bachrach, Y., Key, P., Graepel, T.: Quality expectation-variance tradeoffs in crowdsourcing contests. In: AAAI 2012, Toronto, Canada (2012)
29. Gebrehiwot, M.E., Aalto, S., Lassila, P.: Optimal sleep-state control of energy-aware M/G/1 queues. In: Proceedings of the 8th International Conference on Performance Evaluation Methodologies and Tools (Value tools), Bratislava, Slovakia (2014)
30. Grbac, T.G., Huljenić, D.: On the probability distribution of faults in complex software systems. Inf. Softw. Technol. **58**, 250–258 (2016)
31. Ghodsi, A., Godfrey, P.B., McKeown, N., Parulkar, G., Raghavan, B.: Architecting for innovation. ACM Comput. Commun. Rev. **41**(3), 24–36 (2011)
32. Guo, T., Sharma, U., Wood, T., Sahu, S., Shenoy, P.: Seagull: intelligent cloud bursting for enterprise applications. In: 2012 USENIX Annual Technical Conference (USENIX ATC 2012), USENIX Association, Berkeley, CA, USA (2012)

33. Gustarini, M., Scipioni, M., Fanourakis, M., Wac, K.: Differences in smartphone usage: validating, evaluating, and predicting mobile user intimacy. Pervasive Mob. Comput. **30**, 50–72 (2016)
34. Gwak, Y., Kim, R.Y.: A novel wireless TCP for 5G mobile networks. World J. Wirel. Devices Eng. **1**(1), 1–6 (2017)
35. Hakiri, A., et al.: Software-defined networking: challenges and research opportunities for future internet. Comput. Netw. **75**, 453–471 (2014)
36. Hatton, L.L.: Power-law distributions of component size in general software systems. IEEE Trans. Softw. Eng. **35**(4), 566–572 (2009)
37. Hoßfeld, T., Schatz, R., Biersack, E., Plissonneau, L.: Internet video delivery in youtube: from traffic measurements to quality of experience. In: Biersack, E., Callegari, C., Matijasevic, M. (eds.) Data Traffic Monitoring and Analysis. LNCS, vol. 7754, pp. 264–301. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36784-7_11
38. Hoßfeld, T., Schatz, R., Varela, M., Timmerer, C.: Challenges of QoE management for cloud applications. IEEE Commun. Mag. **50**(4), 28–36 (2012)
39. Hoßfeld, T., Skorin-Kapov, L., Haddad, Y., Pocta, P., Siris, V., Zgank, A., Melvin, H.: Can context monitoring improve QoE? a case study of video flash crowds in the internet of services. In: IFIP/IEEE International Symposium on Integrated Network Management (IM), pp. 1274–1277 (2015)
40. Hyytiä, E., Righter, R., Aalto, S.: Task assignment in a server farm with switching delays and general energy-aware cost structure. Perform. Eval. **75–76**, 17–35 (2014)
41. Hyytiä, E, Righter, R., Aalto, S.: Energy-aware job assignment in server farms with setup delays under LCFS and PS. In: Proceedings of the 26th International Teletraffic Congress (ITC'26), Karlskrona, Sweden (2014)
42. Ickin, S., et al.: Factors influencing quality of experience of commonly used mobile applications. IEEE Commun. Mag. **50**(4), 48–56 (2012)
43. IEEE 802.1 Time-Sensitive Networking Task Group. http://www.ieee802.org/1/pages/tsn.html
44. Jain, S., Kumar, A., Mandal, S., et al.: B4: experience with a globally-deployed software defined WAN. In: SIGCOMM 2013. ACM, New York (2013)
45. Jakab, J., Ganchev, I., O'Droma, M.: Third-party charging and billing for the ubiquitous consumer wireless world. Int. J. Commun. Ant. Prop. **1**(2), 136–144 (2011)
46. Jarschel, M., et al.: On the accuracy of leveraging SDN for passive network measurements. In: ATNAC 2013, Christchurch, New Zealand (2013)
47. Jatoth, C., Gangadharan, G.R., Buyya, R.: Computational intelligence based QoS-aware web service composition: a systematic literature review. IEEE Trans. Serv. Comput. **10**(3), 475–492 (2017)
48. Ji, Z., Ganchev, I., O'Droma, M.: An IWBC consumer application for always best connected and best served: design and implementation. IEEE Trans. Consum. Electron. **57**(2), 462–470 (2011)
49. Jones, T., Fairhurst, G., Perkins, C.: Raising the datagram API to support transport protocol evolution. In: Proceedings of the 16th International IFIP TC6 Networking Conference (Networking 2017), Stockholm (2017)
50. Kairi, A., Chakraborty, S.: A study of packet control techniques for wireless network. Am. J. Adv. Comput. **3**(1), 38–45 (2016)
51. Kash, I., Key, P.: Pricing the cloud. IEEE Internet Comput. **20**, 36–43 (2016)
52. Khurshid, A., Kabir, M.H., Prodhan, M.A.T.: An improved TCP congestion control algorithm for wireless networks. In: Proceedings of the 13th IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, August 2011, pp. 382–387 (2011)

53. Kreutz, D., et al.: Software-defined networking: a comprehensive survey (2014). http://arxiv.org/abs/1406.0440
54. Laghari, K.U.R., Crespi, N., Connelly, K.: Toward total quality of experience: a QoE model in a communication ecosystem. Commun. Mag. IEEE **50**(4), 58–65 (2012)
55. Le Callet, P., Möller, S., Perkis, A., eds.: Qualinet white paper on definitions of quality of experience. European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Lausanne, Switzerland, v1.2 (2012)
56. Leitner, P., Hummer, W., Dustdar, S.: Cost-based optimization of service compositions. IEEE Trans. Serv. Comput. **6**(2), 239–251 (2013)
57. Leitner, P., Zabolotnyi, R., Gambi, A., Dustdar, S.: A framework and middleware for application-level cloud bursting on top of infrastructure-as-a-service clouds. In: 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing (UCC 2013), pp. 163–170. IEEE Computer Society, Washington (2013)
58. Lemos, A.L., Daniel, F., Benatallah, B.: Web service composition: a survey of techniques and tools. ACM Comput. Surv. **48**(3), Art. 33, 1–41 (2015)
59. Li, M.: Cross-layer resource control to improve TCP performance over wireless network. In: Proceedings of the 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007), pp. 706–711 (2007)
60. Liberal, F., Taboada, I., Fajardo, J.O.: Dealing with energy-QoE trade-offs in mobile video. J Comput. Netw. Commun. 2013 (2013). Article ID 412491, 12 pages
61. Lu, F., Du, H., Jain, A., Voelker, G.M., Snoeren, A.C., Terzis, A.: CQIC: revisiting cross-layer congestion control for cellular networks. In: Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (HotMobile 2015), pp. 45–50 (2015)
62. McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J.: OpenFlow: enabling innovation in campus networks. SIGCOMM Comput. Commun. Rev. **38**(2), 69–74 (2008)
63. Mell, P., Grance, T.: The NIST definition of cloud computing. National Institute of Standards and Technology, NIST Special Publication 800-145 (2011)
64. Metzger, F., Liotou, E., Moldovan, C., Hoßfeld, T.: TCP video streaming and mobile networks: not a love story, but better with context. Comput. Netw. **109**(2016), 246–256 (2016)
65. O'Droma, M., Ganchev, I.: The creation of a ubiquitous consumer wireless world through strategic ITU-T standardization. IEEE Commun. Mag. **48**(10), 158–165 (2010)
66. Pathak, N.: A look at TCP Initial congestion window in content delivery networks (2017). https://blog.imaginea.com/look-at-tcp-initcwnd-cdns. Accessed 12 Oct 2017
67. Raiss-el-Fenni, M., et al.: POMDP game framework for service providers inciting mobile users. In: IEEE International Conference on Communication (ICC2014), Sydney, Australia (2014)
68. Reichl, P., Egger, S., Möller, S., Kilkki, K., Fiedler, M., Hossfeld, T., Tsiaras, C., Asrese, A.: Towards a comprehensive framework for QoE and user behavior modelling. In: Proceedings of the 7th International Workshop on Quality of Multimedia Experience (QoMEX) (2015)
69. Reiter, U., Brunnström, K., De Moor, K., Larabi, M.C., Pereira, M., Pinheiro, A., You, J., Zgank, A.: Factors influencing quality of experience. In: Quality of Experience, pp. 55–72 (2014)
70. Scheuner, J., Leitner, P., Cito, J., Gall, H.: Cloud WorkBench – infrastructure-as-code based cloud benchmarking. In: Proceedings of the 6th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2014) (2014)

71. Shannon, J.O., Flaithearta, P., Cinar, Y., Melvin, H.: Enhancing multimedia QoE via more effective time synchronisation over 802.11 Networks. In: Proceedings of the 7th International Conference on Multimedia Systems ACM MMSys 2016, Article No. 25 (2016)
72. Srivastava, A.: Performance analysis of QUIC protocol under network congestion. Master Thesis, Worcester Polytechnic Institute (2017)
73. TAACCS: Cross-Disciplinary Research on Time-Aware Applications, Computers, and Communication Systems. http://www.taaccs.org/index.html
74. Taboada, I., Fajardo, J.O., Liberal, F.: QoE and energy-awareness for multi-layer video broadcasting. In: IEEE International Conference on Multimedia and Expo (ICME2011), Barcelona, Spain (2011)
75. Taboada, I., Liberal, F.: A novel scheduling index rule for QoE maximization in wireless networks. In: Abstract and Applied Analysis 2014 (2014)
76. Trammell, B., Hildebrand, J.: The trouble with transports – moving UDP to layer 3.5 (2014). http://www.ietf.org/proceedings/90/slides/slides-90-appsawg-7.pdf
77. Valerio, V.D., Cardellini, V., Lo Presti, F.: Optimal pricing and service provisioning strategies in cloud systems: a Stackelberg game approach. In: IEEE 6th International Conference on Cloud Computing (CLOUD 2013), Santa Clara, CA, pp. 115–122 (2013)
78. Wac, K., Fiordelli, M., Gustarini, M., Rivas, H.: Quality of life technologies: experiences from the field and key research challenges. IEEE Internet Comput. 19(4), 28–35 (2015)
79. Wac, K., et al.: mQoL: experimental methodology for longitudinal, continuous quality of life assessment via unobtrusive, context-rich mobile computing in situ. In: The International Society for Quality-of-Life Studies Conference (ISQOLS 2016) (2016)
80. Wang, Z., Zeng, X., Liu, X., Xu, M., Wen, Y., Chen, L.: TCP congestion control algorithm for heterogeneous internet. J. Netw. Comput. Appl. 68(C), 56–64 (2016)
81. Weiss, M., Melvin, H., Chandoke, S.: Time: an integral part of cyber-physical systems. In: Time-Sensitive Networks and Applications (TSNA) (2016)
82. White, G.: Analysis of Google SPDY and TCP Initcwnd. Cable Labs (2012). https://www.cablelabs.com/wp–content/uploads/2014/05/Analysis_of_Google_SPDY_TCP.pdf. Accessed 12 Oct 2017
83. Wilhelm, R., Grund, D.: Computation takes time, but how much? Commun. ACM 57(2), 94–103 (2014)
84. Yin, Q., Kaur, J., Smith, D.: TCP rapid: from theory to practice. Technical report 17-001, Department of Computer Science, UNC Chapel Hill (2017)
85. Zeng, L., Benatallah, B., Ngu, A.H.H., Dumas, M., Kalagnanam, J., Chang, H.: QoS-aware middleware for web services composition. IEEE Trans. Soft. Eng. 30(5), 311–327 (2004)