A Survey on Service-Oriented Network Virtualization Toward Convergence of Networking and Cloud Computing

Qiang Duan, Yuhong Yan, and Athanasios V. Vasilakos

Abstract—**The crucial role that networking plays in Cloud computing calls for a holistic vision that allows combined control, management, and optimization of both networking and computing resources in a Cloud environment, which leads to a convergence of networking and Cloud computing. Network virtualization is being adopted in both telecommunications and the Internet as a key attribute for the next generation networking. Virtualization, as a potential enabler of profound changes in both communications and computing domains, is expected to bridge the gap between these two fields. Service-Oriented Architecture (SOA), when applied in network virtualization, enables a Network-as-a-Service (NaaS) paradigm that may greatly facilitate the convergence of networking and Cloud computing. Recently the application of SOA in network virtualization has attracted extensive interest from both academia and industry. Although numerous relevant research works have been published, they are currently scattered across multiple fields in the literature, including telecommunications, computer networking, Web services, and Cloud computing.**

In this article we present a comprehensive survey on the latest developments in service-oriented network virtualization for supporting Cloud computing, particularly from a perspective of network and Cloud convergence through NaaS. Specifically, we first introduce the SOA principle and review recent research progress on applying SOA to support network virtualization in both telecommunications and the Internet. Then we present a framework of network-Cloud convergence based on service-oriented network virtualization and give a survey on key technologies for realizing NaaS, mainly focusing on state of the art of network service description, discovery, and composition. We also discuss the challenges brought in by network-Cloud convergence to these technologies and research opportunities available in these areas, with a hope to arouse the research community's interest in this emerging interdisciplinary field.

Index Terms—**Network virtualization, the service-oriented architecture, cloud computing, network-as-a-Service (NaaS).**

I. Introduction

ONE of the most significant recent advances in the field of information technology is Cloud computing. Cloud computing is a large scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically scalable computing functions and services are delivered on demand to external customers over the Internet [1]. A technical foundation of Cloud computing lies in the virtualization of various computing resources, which is essentially an abstraction of logical functions from underlying physical resources.

Networking plays a crucial role in Cloud computing. Cloud services normally represent remote delivery of computing resources, most often via the Internet. This is especially relevant in public Cloud environments where customers obtain services from a third-party Cloud provider. From a service provisioning perspective, Cloud services consist of not only computing functions provided by Cloud infrastructure but also communications functions offered by the Internet. Networking is also a key element for providing data communications in Cloud data centers as well as among data centers distributed at different locations. Results obtained from recent study on Cloud computing performance have indicated that networking performance has a significant impact on the quality of Cloud services, and in many cases data communications become a bottleneck that limits Clouds from supporting high-performance applications [2] [3]. Therefore, networks with Quality of Service (QoS) capabilities become an indispensable ingredient for high-performance Cloud computing.

For example, a high performance application utilizes the Cloud infrastructure for storing and processing a large set of data with a requirement on the maximum service response delay (the time period that the application has to wait for receiving results back from the Cloud after it starts transmitting data to the Cloud). This application may use the storage capacity of Amazon S3 (Simple Storage Service) and the computing capability provided by Amazon EC2 (Elastic Compute Cloud). In order to make the Cloud services available to the application, the underlying network infrastructure must also provide network services for transmitting data from the application to the S3 virtual disk, supporting data communications between the virtual disk and the EC2 virtual machine (Amazon EC2 and S3 servers may be located at different geographical locations that are connected via the Internet), and delivering processing results back to the application. Therefore, the service offered to the Cloud user (this application) is essentially a composition of both Cloud and network services. In order to meet the service delay requirement of the application, sufficient amount of networking resources (e.g. transmission bandwidth and packet forwarding capacity)

Manuscript received June 14, 2012; revised September 24, 2012. The associate editor coordinating the review of this paper and approving it for publication was H. Lutfiyya.

Q. Duan is with the Information Science & Technology Department, Pennsylvania State University Abington College (e-mail: qduan@psu.edu).

Y. Yan is with the Department of Computer Science & Software Engineering, Concordia University (e-mail: yuhong@cse.concordia.ca).

A. V. Vasilakos is with the Computer Science Department, Kuwait University (e-mail: vasilako@ath.forthnet.gr, vasilakos@sci.kuniv.edu.kw).

Digital Object Identifier 10.1109/TNSM.2012.113012.120310

must be provided to guarantee network delay performance in addition to the computing and storage resources offered by the Cloud infrastructure for meeting data processing and storing requirements.

The significant role that networking plays in Cloud computing calls for a holistic vision of both computing and networking resources in a Cloud environment. Such a vision requires underlying network infrastructure to be opened and exposed to upper layer applications in Clouds; thus enabling combined management, control, and optimization of computing and networking resources for Cloud service provisioning. This leads to a convergence of networking and Cloud computing systems toward a composite network-Cloud service provisioning system. Due to the complexity of networking technologies and protocols, exposure of network functionalities in a Cloud environment is only feasible with appropriate abstraction and virtualization of networking resources.

On the other hand, telecommunication and networking systems are facing the challenge of rapidly developing and deploying new functions and services for supporting the diverse requirements of various computing applications [4]. Fundamental changes are required in the Internet architecture to allow heterogeneous networking systems to coexist and cooperate for supporting a wide spectrum of applications. A promising approach that the networking research community takes for addressing these challenges is virtualization of networking resources; namely decoupling service provisioning from network infrastructure and exposing underlying network functionalities through resource abstraction and virtualization. Such an approach in general is described by the term *network virtualization*, which is expected to be a key attribute of the future networking paradigm [5].

As a potential enabler of profound changes in both computing and communications domains, virtualization is expected to bridge the gap between these two fields that traditionally live quite apart and enable a convergence of networking and Cloud computing. Network virtualization in a Cloud environment allows a holistic vision of both computing and networking resources as a single collection of virtualized, dynamically provisioned resources for composite network-Cloud service provisioning. Convergence of networking and Cloud computing is likely to open up an immense field of opportunities to the IT industry and allows the next generation Internet to provide not only communication functions but also various computing services. Various telecommunications and Internet service providers around the world have already shown a great deal of interest in providing Cloud services based on their network infrastructure. For example, AT&T has launched its Cloud architecture that offers a wide range of enterprise hosting and Cloud computing services[1]. Verizon has also started offering enterprise Cloud computing services based on its Verizon Cloud Platform[2].

Convergence of networking and Cloud computing can be viewed from vertical and horizontal dimensions. In the vertical dimension, resources and functionalities in network infrastructure are opened and exposed through an abstract virtualization interface to upper layer functions in the Cloud, including resource management modules and other functions for offering Cloud services. In the horizontal dimension, Cloud data centers that offer computing functions and network infrastructure that provide data communications converge into a composite network-Cloud service provisioning system. In both dimensions, such a convergence enables combined management, control, and optimization of networking as well as computing resources in a Cloud environment.

Some technical issues must be addressed for realizing the notion of convergence between networking and Cloud computing. Key requirements for network-Cloud convergence include networking resource abstraction and exposure to upper layer applications and collaborations among heterogeneous systems across the networking and computing domains. Therefore an important research problem is to develop the mechanism for supporting effective, flexible, and scalable interaction among key players in a converged networking and Cloud computing environment, including networking and computing infrastructure providers, networking and computing service providers, and various applications as the customers of composite network-Cloud services. Service-Oriented Architecture (SOA), when applied in both network virtualization and Cloud computing, offers a promising approach to enabling the network-Cloud convergence.

SOA provides effective architectural principles for heterogeneous system integration. Essentially service-orientation facilitates virtualization of computing systems by encapsulating system resources and capabilities in the form of services and provides a loose-coupling interaction mechanism among these services [6]. SOA has been widely adopted in Cloud computing via the paradigms of Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). Applying SOA in the field of networking supports encapsulation and virtualization of networking resources in the form of SOA-compliant *network services*. Service-oriented network virtualization enables a *Network-as-a-Service* (*NaaS*) paradigm that allows network infrastructure to be exposed and utilized as network services, which can be composed with computing services in a Cloud environment. Therefore the NaaS paradigm may greatly facilitate a convergence of networking and Cloud computing.

Currently Web services provide the main implementation approach for SOA. Web service technologies, including service description, discovery, and composition are mainly developed in the distributed computing area; therefore evolution to these technologies is needed to meet the requirements of NaaS toward network-Cloud convergence. Application of SOA in networking recently formed an active research area that has attracted extensive attention from both industry and academia. A great amount of research efforts have been made on key technologies for NaaS, including network service description, discovery, and composition. These works are conducted in various fields scattered across telecommunications, computer networking, Web services, and distributed computing. Although some relevant surveys have been published, for example [4], [7], and [8], they each focus on a particular field, either telecommunication services or Web services/distributed computing systems. In addition, they all lack discussion on

[1]http://cloudarchitect.att.com/Home/
[2]http://www.verizonbusiness.com/solutions/bysolutions/cloud/

integrating network services in a Cloud computing environment. This motivates a comprehensive survey in the literature that reflects the current status of service-oriented network virtualization for network-Cloud convergence, which is the main objective of this article.

In this article, we first introduce the SOA concept and principle and examine the latest developments of SOA-based virtualization in both telecommunications and the Internet. Then we discuss convergence of networking and Cloud computing and present a framework of service-oriented network virtualization for network-Cloud convergence. We also give a comprehensive survey on key enabling technologies of the NaaS paradigm for supporting network-Cloud convergence, mainly focusing on network service description, discovery, and composition technologies. Challenges brought in by network-Cloud convergence to these technologies and opportunities for future research in these areas are also discussed in this article. We hope to arouse interest of the research community on this emerging interdisciplinary field, where cross-fertilization among multiple areas may yield innovative solutions that will significantly enhance performance of the future Cloud-based information infrastructure.

II. The Service-Oriented Architecture

The service-orientation principle means that the logic required to solve a large problem can be better constructed, carried out, and managed, if it is decomposed into a collection of smaller and related pieces, each of which addresses a concern or a specific part of the problem. Service-Oriented Architecture (SOA) encourages individual units of logic to exist autonomously yet not isolated from each other. Within SOA, these units are known as services [6].

SOA provides an effective solution to coordinating computational resources across heterogeneous systems to support various application requirements. As described in [9], SOA is an architecture within which all functions are defined as independent services with invokable interfaces that can be called in defined sequences to form business processes. SOA can be considered as a philosophy or paradigm for organizing and utilizing services and capabilities that may be under the control of different ownership domains [10]. Essentially SOA enables virtualization of various computing resources in form of services and provides a flexible interaction mechanism among services.

A service in SOA is a module that is self-contained (i.e., the service maintains its own states) and platform-independent (i.e., interface to the service is independent of its implementation platform). Services can be described, published, located, orchestrated, and programmed through standard interfaces and messaging protocols. All services in SOA are independent of each other and service operations are perceived as opaque by external services, which guarantees that external components neither know nor care how services perform their functions. The technologies providing the desired functionality of the service are hidden behind the service interface.

A key feature of SOA is loosely-coupled interaction among heterogeneous systems in the architecture. The term "coupling" indicates the degree of dependency any two systems have on each other. In loosely coupled interaction, systems need not know how their partners behave or are implemented, which allows systems to connect and interact more freely. Therefore, loose-coupling of heterogeneous systems provides a level of flexibility and interoperability that cannot be matched using traditional approaches for building highly integrated, cross-platform, inter-domain communication environments. Other features of SOA include reusable services, formal contract among services, service abstraction, service autonomy, service discoverability, and service composability. These features make SOA a very effective architecture for heterogeneous system integration with resource virtualization to support diverse application requirements.

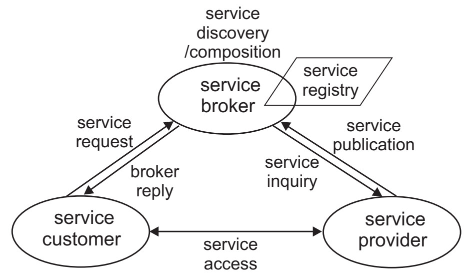

Fig. 1. A Web services-based SOA implementation.

Though SOA can be implemented with different technologies, Web services provide a preferred environment for realizing SOA. A Web service has an interface that describes a collection of operations that are network accessible through standardized XML messaging [11]. A Web service is described using a standard, formal XML notion, called its service description. It covers all the details necessary to interact with the service, including message formats, transport protocols, and location. The interface hides the implementation details of the service, allowing it to be used independently of its implementation. This enables Web services-based systems to be loosely coupled, component-oriented, with cross-technology implementations.

Key elements of a Web service-based implementation of SOA include service provider, service broker/registry, and service customer. The basic operations involved in the interaction among these elements are service description publication, service discovery, and service binding/access. In addition, service composition is also an important operation for meeting customers' service requirements. The key Web service elements and operations are shown in Fig. 1. A service provider makes its service available in the system by publishing a service description at a service registry. Service discovery, typically performed by a broker, is the process that responds to a customer request for discovering a service that meets specified criteria. Multiple services may be composed into a composite service to meet the customer's requirements.

Representational State Transfer (REST) can be considered as an alternative to the standard Web service technologies for implementing SOA systems [12]. In REST, the focus is on the resources exposed by services. Each resource is identified by a URI represented by a certain MIME type (such as XML or JSON), and accessed and controlled using POST, GET, PUT, or DELETE http methods. SOA services based on REST expose the resources that they manage through (and only through) these four methods. This set of technologies that follow the REST design style for realizing service-orientation is conventionally called RESTful Web services.

The SOA principle and its implementation technologies have become state of the art of information service delivery and have been widely applied in various distributed computing areas, including Cloud computing. SOA enables more flexible and reusable services that may be reconfigured and augmented more swiftly than traditional system construction; thus can accelerate time-to-business objective and result in better business agility. SOA also provides a standard way to represent and interact with application functionalities thus improving interoperability and integration across heterogeneous systems.

III. Service-Oriented Network Virtualization in Telecommunications

An important aspect of telecommunications evolution in the past decades has been to create new market-driven applications by reusing existing service components. The methodology taken by the telecom research and development community for achieving this objective is based on the idea of separating service-related functions from data transport mechanisms. Such separation allows underlying network infrastructure to be virtualized and shared by service-related functions in order to create various applications. This is essentially the notion of virtualization in the telecommunications domain. In recent years, the SOA principle and Web service technologies have been applied to facilitate virtualization in telecom systems.

Early efforts toward making telecom network a programmable environment for delivering value-added services can be traced back to Intelligent Network (IN) [13]. The IN idea is to define overlay service architecture on top of physical network infrastructure and extract service intelligence into dedicated service control points. Later on some telecom API standards, including Parlay, Open Service Architecture (OSA), and Java API for Integrated Networks (JAIN), were developed for achieving a similar objective as IN but with easier service development [4]. These APIs simplified telecom service development by abstracting signaling protocol details of the underlying networks.

Although these technologies were promising, they lacked an effective mechanism for realizing the separation of service provisioning and network infrastructure. Remote procedure call and functional programming conceptually drove the IN realization. Parlay/OSA and JAIN were typically implemented based on Common Object Request Broker Architecture (CORBA) and Java Remote-Method Invocation (RMI) technologies. System modules in such distributed computing technologies are essentially tightly coupled; therefore they lack full support for networking resource abstraction and virtualization.

In early 2000s a simplified version of the Parlay/OSA API called Parlay X was developed jointly by the Parlay Group, ETSI, and 3GPP [14]. Parlay X is based on the emergence of Web service technologies. The objective of Parlay X is to offer a higher level of abstraction than Parlay/OSA in order to allow developers to design and build value-added telecom applications without knowing details of networking protocols. Web service technologies are employed in Parlay X to expose networking capabilities to upper layer applications, which opens a door for applying SOA to realize the separation of service provisioning and data transportation.

Telecom systems are undergoing a fundamental transition toward a multi-service packet-switching IP-based network. Two representative developments in the transition are the Next Generation Network (NGN) [15] and IP-based Multimedia Subsystem (IMS) [16]. NGN is defined by ITU-T as a packet-based network able to provide services, including telecom services, and able to make use of multiple broadband, QoS-enabled transport technologies. In NGN, service-related functions are independent from underlying transport technologies. IMS is an effort made by telecom-oriented standard bodies, such as 3GPP and ETSI, to realize the NGN concept that presents an evolution from the traditional closed signaling system to the NGN service control system [17]. ITU-T has also developed a specification for NGN service integration and delivery environment [18]. A key feature of the NGN architecture is the decoupling of network transport and service-related functions, which allows virtualization of network infrastructure for flexible service provisioning.

Recently rapid development of new services became a crucial requirement to telecom operators. However, telecom systems have been designed specifically to support a narrow range of precisely defined communication services, which are implemented on fairly rigid infrastructure with minimal capabilities for ad hoc reconfiguration. Operation and management functions in traditional networks are also specifically designed and customized to facilitate particular types of services. Tightly coupling between service provisioning and network infrastructure becomes a barrier to rapid and flexible service development and deployment. In order to resolve the problem of "silo" mode of telecom service provisioning, research and development efforts have been made for building a Service Delivery Platform (SDP). At a high level, SDP is a framework that facilitates and optimizes all aspects of service delivery including service design, development, provisioning, and management. The core idea is to have a framework for service management and operation by aggregating the network capabilities and service management functions in a common platform. Main SDP specifications include OMA Open Service Environment (OSE) [19] and TM Forum Service Delivery Framework (SDF) [20].

The objective of SDP is to provide an environment in which upper layer applications can be easily developed by combining underlying networking capabilities and also enable collaboration across network service providers, content providers, and third-party service providers. The virtualization concept and SOA principle play a key role in both OSE and SDF specifications to achieve this objective. The method taken by both specifications is to define a set of standard service components called service enablers and develop a framework that allows new services to be built by composing service enablers. The service enablers support virtualization of networking resources by encapsulating underlying network functionalities through a standard abstract interface. The web services approach has become a de facto standard for communications among system components in SDP. Web service orchestration technologies such as BPEL [21] are also becoming part of SDP for enabling services to be composed with both telecom functional blocks and business logic/applications in the computing domain.

As users consume services offered through various networks, they have pushed for blending the service offerings of various providers for a richer experience. In order to allow network and computing service providers, content providers, and end-users to offer and consume collaborative services, there is a need for an efficient way of service and application delivery that at the same time is customer-centric. This is a challenge that has not been sufficiently addressed by the aforementioned developments such as IMS, NGN, and SDP. Therefore, there has been a motivation to organize the services/applications offered by various networks on an overlay that allows service providers to offer rich services. Toward this objective, IEEE recently developed the Next Generation Service Overlay Network (NGSON) standard [22]. NGSON specifies context-aware, dynamically adaptive, and self-organizing networking capabilities including both service level and transport level functions that are independent of the underlying network infrastructure. NGSON aims to bridge the service layer and transport network over IP infrastructure to address the accommodation of highly integrated services. NGSON particularly focuses on composing new collaborative services by either using existing components (from IMS, NGN, SDP, etc.) or defining new NGSON components. Key functional entities of NGSON service control include service discovery and negotiation, service routing, and service composition. Web service technologies have been employed to realize these functional entities [23], [24], [25].

The aforementioned review shows that recent evolution of service management in telecommunications has followed a path toward network virtualization; that is, decoupling service provisioning from data transport and exposing network infrastructure through resource abstraction. The SOA principle and Web service technologies play a key role in facilitating virtualization in telecom systems.

IV. Service-Oriented Network Virtualization in the Future Internet

Fundamental changes in network architecture and service delivery model are required by the future Internet. However, the current IP-based Internet protocol along with the huge amount of investment in the Internet infrastructure make any disruptive innovation in the Internet architecture very difficult. In order to fend off the ossification of the current Internet, network virtualization has been proposed as a key attribute of the future inter-networking paradigm. Network virtualization in the Internet can be described as a networking environment that allows one or multiple service providers to compose heterogeneous virtual networks that co-exist together but in isolation from each other and to deploy customized end-to-end services on those virtual networks by effectively sharing and utilizing underlying network resources provided by infrastructure providers [5].

Essentially network virtualization follows a well-tested principle - separation of policy from mechanism - in the Internet.

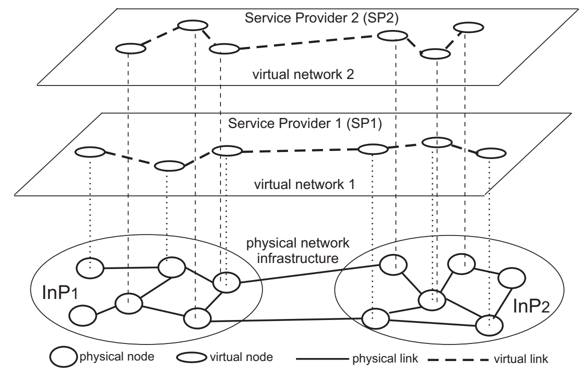

Fig. 2. Illustration of a network virtualization environment.

In this case, network service provisioning is separated from data transportation mechanisms; thus dividing the traditional role of Internet service providers into two entities: Infrastructure Providers (InPs) who manage the physical infrastructure and network Service Providers (SPs) who create virtual networks for offering end-to-end network services by utilizing resources obtained from InPs. Key attributes of network virtualization include abstraction (details of the network resources are hidden), indirection (indirect access to network resources that may be combined to form different virtual networks), resource sharing (network elements can be partitioned and utilized by multiple virtual networks), and isolation (loose or strict isolation between virtual networks). Physical network infrastructure, consisting of links and nodes, is virtualized and made available to virtual networks, which can be setup and torn down dynamically by SPs according to customer needs. Fig. 2 illustrates a network virtualization environment, in which the service providers SP1 and SP2 construct two virtual networks by using resources obtained from the infrastructure providers InP1 and InP2.

Network virtualization has a significant impact on the next generation networking. Allowing multiple virtual networks to cohabit on shared physical infrastructure, network virtualization provides flexibility, promotes diversity, and promises increased manageability. The best-effort Internet today is basically a commodity service that gives network service providers limited opportunities to distinguish themselves from competitors. A diversified Internet enabled by network virtualization offers a rich environment for innovations; thus stimulating development and deployment of new Internet services. Network virtualization enables a single SP to obtain control over the entire end-to-end service delivery path across network infrastructure that may belong to different domains, which will greatly facilitate the end-to-end QoS provisioning.

Network virtualization has attracted extensive research interest from both academia and industry. Virtualization was first employed in the Internet as an approach to developing virtual test beds for new network architecture and protocols, for example in the PlanetLab [26] and GENI [27] projects. Then the role of virtualization in the Internet has evolved from a research method to a fundamental attribute of the inter-networking paradigm [28]. CABO proposed in [29] is new Internet architecture that decouples network service providers and infrastructure providers to support virtual networks over a shared physical substrate. 4WARD is a large EU FP7 project in which network virtualization is employed as a key technology to allow virtual networks to operate in parallel in future Internet [30]. FEDERICA is another FP7 project with a core objective to create a Europe-wide infrastructure of network resources that can be sliced to provide a virtual Internet environment [31]. The concept of network virtualization is also employed in the AGAVE project for developing an open end-to-end Internet service provisioning solution [32]. The line of research on Software Defined Network (SDN), for example the OpenFlow protocol that is currently under active study, also follows the virtualization principle by separating network control from the data plane [33]. More relevant works on network virtualization can be found in the survey [5].

Some standard organizations have also started working on network virtualization specifications. For example, in July 2009 ITU-T established the Focus Group on Future Network (FG-FN), in which network virtualization is identified as one of the fundamental study topics. The Internet Research Task Force (IRTF) created the Virtual Networks Research Group (VNRG) in early 2010, which specifically focuses on network virtualization.

The recent developments in telecommunications discussed in the previous section, such as Parlay X, NGN, SDP, and NGSON, are all based on the principle of separating service provisioning from network infrastructure, which is essentially virtualization in the telecom domain. Comparison between virtualization in telecommunications and in the Internet shows some difference in the perspective and emphasis of the telecom and Internet communities regarding embracing the notion of virtualization in their domains. Virtualization in telecommunication systems tends to focus on exposing networking platform to upper layer applications for facilitating rapid development of value-added services. Therefore virtualization is typically realized above the network layer through standard APIs with abstraction of networking resources and functionalities. The Internet community aims at adopting virtualization as a key attribute of the core network architecture, therefore tends to realize virtualization on or below the network layer. In addition, Internet virtualization has been developed with an important objective to enable heterogeneous network architecture, including both IP-based and non IP-based architecture, to coexist and cooperate in the future Internet. Though virtualization in telecom systems in principle supports a service delivery platform that is independent of the underlying networking technologies, most of the current specifications, for example NGN, IMS, and SDP, assume IP-based packet switching architecture for the physical network infrastructure.

Although significant progress has been made toward network virtualization, there are still many challenges that must be addressed before this notion can be fully realized in the future Internet. In order to create and provision virtual networks for meeting users' requirements, SPs first need to discover available resources in network infrastructure that may belong to multiple administrative domains; then the appropriate network resources need to be selected and composed to form virtual networks. Therefore a key to realize the network virtualization lies in flexible and effective interaction and

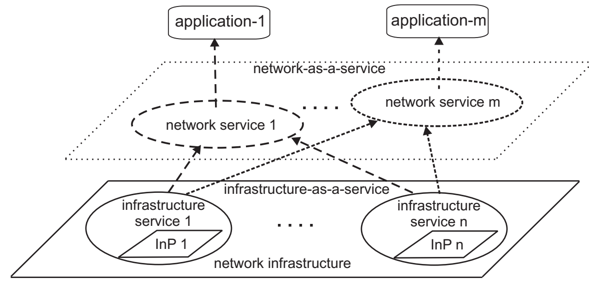

Fig. 3. Service-oriented network virtualization.

collaboration among InPs, SPs, and applications (as end users of virtual networks). SOA, as a very effective architecture for heterogeneous system integration, offers a promising approach to facilitating network virtualization in the future Internet. A layered structure for service-oriented network virtualization is shown in Fig. 3. Following the SOA principle, resources in network infrastructure can be encapsulated into *network infrastructure services*. SPs access networking resources via the Infrastructure-as-a-Service paradigm and compose the infrastructure services into end-to-end *network services*. Applications, as the end users of virtual networks, utilize the underlying networking platform by accessing the network services offered by SPs, which is essentially a *Network-as-a-Service* (*NaaS*) paradigm.

Service-oriented network virtualization has become an active research area that attracts extensive interest. In UCLPv2 (User Controlled Light Path), a Canadian research project for enabling user control and management of optical network infrastructure, Web service technologies were employed to expose resources in optical network infrastructure as services [34]. The framework of network infrastructure service developed in UCLPv2 then evolved into a number of different projects, including Argia as a commercial implementation for optical network infrastructure services, Ether for developing Ethernet and MPLS infrastructure services, and MANTICORE for supporting logical IP network as services [35]. In [36] the authors designed a transport stratum according to the SOA paradigm in order to expose transport functionalities as services to the service stratum in NGN. A service-oriented network virtualization architecture was developed in [37], which consists of physical infrastructure layer, virtual network layer, and service network layer from bottom to top. Analytical modeling and analysis techniques for evaluating end-to-end QoS in service-oriented network virtualization have also been developed in [38] and [39]. Service-oriented network virtualization has also been adopted by industry in various networking equipment and solution developments. For example, the Service-Oriented Network Architecture (SONA) [40] developed by Cisco provides a framework for implementing the infrastructure-as-a-service strategy in the networking domain.

Applying SOA in network virtualization makes loose-coupling a key feature of interaction among InPs, SPs, and applications. Therefore, the NaaS paradigm inherits the merit of SOA that enables flexible and effective collaboration across heterogeneous networking systems for providing services that meet diverse application requirements. Service-oriented network virtualization also provides a means to present abstracted networking capabilities to upper-layer applications. Because of the heterogeneity of network protocols, equipment, and technologies, exposure of networking capabilities to applications without virtualization would lead to unmanageable complexity. The abstraction of networking resources through service-oriented network virtualization can address the diversity and significantly simplify the interaction between applications and the underlying network platform.

V. Service-Oriented Network Virtualization for Convergence of Networking and Cloud Computing

Virtualization, as a potential enabler of profound changes in both the communications and computing domains, is expected to bridge the gap between these two fields that traditionally live quite apart; thus enabling a convergence of Cloud computing and networking. Network and Cloud convergence allows combined management, control, and optimization of networking as well as computing resources in a Cloud environment. With the advent of virtualized networking technologies, Cloud service delivery could be significantly improved via providing service providers with options to implement virtual networks that offer customized networking solutions for Cloud services.

Serving as a key enabler to virtualization, SOA forms a core element in the technical foundation of Cloud computing. Recent research and development have been bridging the power of SOA and virtualization in the Cloud computing ecosystem [41], [42], [43]. SOA has also been widely adopted as the main model for Cloud service provisioning. Following this model, various virtualized computing resources, including both hardware (e.g. CPU capacity and storage space) and software applications are delivered to customers as services through the Infrastructure-as-a-Service, Platform-as-a-Service, and Software-as-a-Service paradigms. The Open Grid Forum (OGF) is working on the Open Cloud Computing Interface (OCCI) standard [44], which defines SOA-compliant open interfaces for interacting with Cloud infrastructure. Taking a look at some of the most important Cloud providers, we can see that the SOA principle has strongly influenced Cloud service provisioning. For example Amazon, a well-known provider that offers a complete ecosystem of Cloud services including virtual machines (Elastic Compute Cloud EC2) and plain storage (Simple Storage Service S3), exposes its Cloud services via Web service interfaces. GoGrid, another important IaaS provider, defines its interfaces to create and control virtual hardware resources in Cloud infrastructure based on REST technologies, which is an alternative implementation of SOA.

The service-orientation principle, when applied in both network virtualization and Cloud computing, offers a promising approach to facilitating the convergence of networking and Cloud computing. Applying SOA in networking allows virtualization of networking resources in the form of SOA-compliant network services. This enables a Network-as-a-Service (NaaS) paradigm that exposes networking resources and functionalities as services that can be composed with computing services in a Cloud environment. From a service provisioning perspective, the services delivered to end users are essentially composite network-Cloud services that comprise both computing services provided by Cloud infrastructure and network services offered by network infrastructure.

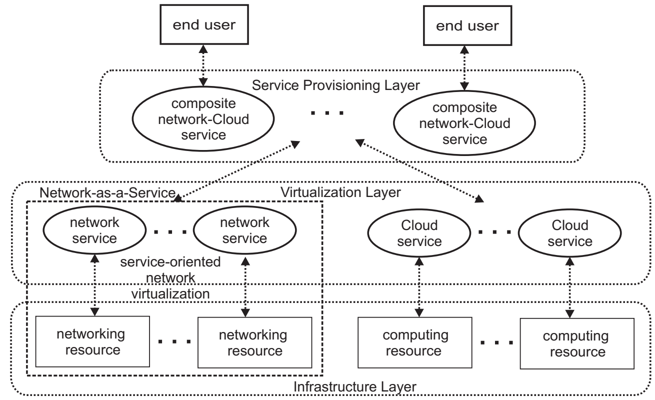

Fig. 4. A NaaS-based framework for network-cloud convergence.

Fig. 4 shows a layered framework for enabling the convergence of networking and Cloud computing via service-oriented network virtualization. In this framework, resources in both networking and computing infrastructure are virtualized into services by following the same SOA principle, which offers a uniform mechanism for coordinating networking and computing systems for Cloud service provisioning. Service-oriented virtualization enables a holistic vision of both networking and computing resources as a single collection of virtualized, dynamically provisioned resources, which allows coordinated management, control, and optimization of resources across the networking and computing domains. In this convergence framework, NaaS enables matching Cloud service requirements with networking capabilities by discovering the appropriate network services. Composition of network and computing services expands the spectrum of Cloud services that can be offered to users. The loose-coupling feature of SOA provides a flexible and effective mechanism in this network-Cloud convergence framework that supports interaction between networking/computing infrastructure and service provisioning functions as well as collaboration among heterogeneous networking and computing systems.

Convergence of networking and Cloud computing has become an important topic in some major research projects. For example, in the IRMOS (Interactive Real-time Multimedia Applications on Service Oriented Infrastructures) project, an Intelligent Service Oriented Network Infrastructure (ISONI) was developed [45]. ISONI consists of a network of resources, including networking, computing, and storage resources, managed and controlled by an ISONI middleware that allows resource sharing among multiple services. The general idea of ISONI is to provide a service-oriented infrastructure for SOA components and services. The NEBULA project sponsored by NSF aims at developing a potential future Internet architecture that provides trustworthy networking for the emerging Cloud computing model of always-available network services [46]. Network virtualization is a core function of the control plane of this new Internet architecture that supports networking and Cloud computing convergence. Cloud networking is an important work package in the EU-funded SAIL (Scalable and

Adaptive Internet Solution) project [47]. The Cloud network architecture developed in this project virtualizes computing, networking, and storage resources as infrastructure services, and particularly addresses network capabilities through the IaaS paradigm to enable composition of computing and network services in a Cloud environment.

Research progress on technologies for SOA-based network-Cloud convergence has also been reported in literature. In [48] the authors proposed a SOA-based Virtual Network Operator (VNO) business model and developed system architecture that uses virtualization to abstract the infrastructure (including networking, computing, and storage) services and compose infrastructure services to meet customer requirements. An architectural solution for future Cloud service provisioning was proposed in [49], which applies SOA in IP network virtualization for supporting Cloud computing. In addition, the concept of marketplace was also employed in this solution to allow trading IP network resources between infrastructure providers and Cloud service providers. An OpenFlow-based network virtualization framework for supporting Cloud computing was proposed in [50]. CoSwitch, a switching mechanism for efficiently support SDN in Cloud data centers, is developed in [51]. The authors of [52] investigated employing NGSON to provide inter-Cloud operations by supporting dynamic routing, composition, and delivery of services through multiple Clouds.

The notion of network-Cloud convergence has also been embraced by the networking industry. Cisco's VFrame Data Center (DC) solution aims to offer service orchestration that enables coordinated provisioning of virtualized networking, computing, and storage resources based on service-oriented network architecture [40]. Cisco also developed the Unified Service Delivery solution with a goal of composing resources in data centers and the end-to-end IP networks for Cloud service provisioning [53].

Standard bodies and industry forums have also started working on related specifications. ITU-T launched a Focus Group on Cloud Computing (FG Cloud) in May 2010, which aims at contributing with the networking aspects for flexible Cloud infrastructure in order to better support Cloud services/applications that make use of communication networks and Internet services [54]. Open Grid Forum (OGF) has formed a work group (NSI-WG) for developing Network Service Interface (NSI) architecture, which encapsulates networking capabilities in form of services that are accessible through a standard interface [55]. A prototype of NSI architecture for standardized global inter-domain provisioning of high performance network connections has been demonstrated in Supercomputing 2011. The Alliance for Telecommunications Industry Solutions (ATIS) launched Cloud Service Forum (CSF) in 2011, which focuses on telecom operators' provision of Cloud services. Its main objective includes exposure of the resources and capabilities of telecom infrastructure as Web services to enable reusing service components provided by different network domains.

As the main approach to implementing SOA, Web services serve as key enabling technologies of the NaaS paradigm that forms a technical foundation for network-Cloud convergence. Fig. 5 gives the high-level structure for a Web service-based delivery system for composite network-Cloud services. In this system, both network and Cloud service providers make their services available by publishing service descriptions at a service registry. When a service consumer, typically an application, needs to utilize a Cloud service, it sends a request to the service broker. The service broker discovers available Cloud and network services by searching the registry and composes the appropriate network and Cloud services into a composite service that meets the consumer's requirements. Please note that Fig. 5 just shows a general service delivery structure, in which service operations including description, discovery, and composition may be realized by various technologies, as surveyed in the rest of this article.

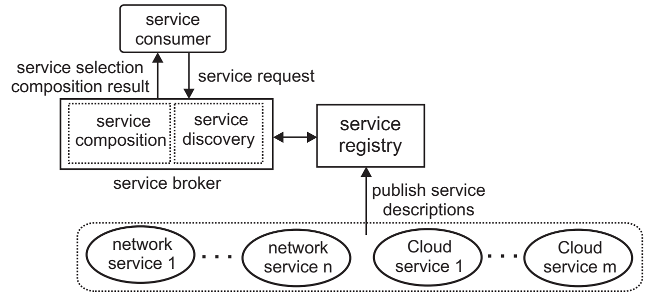

Fig. 5. A Web service-based delivery system for composite network-Cloud services.

For example, in the use case mentioned in Section I where an application utilizes the Cloud for storing and processing data, the application sends the broker a service request that specifies requirements for service functions such as data transmission, processing, and storage, as well as requirements on service performance such as the maximum service delay and the minimum computing capacity and storage space. Typically multiple providers for different types of services exist in a Cloud environment, for example Cloud IaaS providers like Amazon, GoGrid, and Rackspace are available for providing computing capacity and storage space as services; and AT&T, Verizon, and Comcast are available to provide network services for data transmission. The service broker will search service descriptions published at the registry, discover available services, and select a right set of computing, communication, and storage services that can be composed into an end-to-end network-Cloud service for meeting the application requirements. For example, the broker may select Amazon EC2 for meeting the computing capacity requirement and Amazon S3 for meeting the storage space requirement, then select Comcast network service for data transmission between the application host device and S3 virtual disk and Verizon network service for communications between EC2 virtual machine and S3 disk (that are located at different sites). These services work together to meet the requirement on end-to-end service response time, which includes the network delay for data transmission (in both Comcast and Verizon networks), the latency for accessing the S3 virtual disk, and the computing delay introduced by the EC2 virtual machine. Then the broker composes all these participating services into one composite service for the application.

Service-oriented composition of network and Cloud services allows provisioning of network services and Cloud services, which used to be offered separately by different providers, merged into the provisioning of composite network-Cloud services. This convergence enables a new service delivery model in which the roles of traditional Internet service providers and computing service providers merge together into one role of composite network-Cloud service providers. This new service delivery model may stimulate innovations in service development and create a wide variety of new business opportunities. Network-Cloud convergence allows a single-point of visibility of computing and networking operations, providing the opportunity to manage both more effectively. Such a convergence also presents a vital part of an overall cost-saving solution, in which business processes are enhanced and time-to-market is shortened.

As service description, discovery, and composition are key elements of Web services for implementing SOA, description, discovery, and composition of network services are key enabling technologies for the NaaS paradigm that forms a foundation for service-oriented convergence of networking and Cloud computing. The services offered to end users in a Cloud environment are essentially composite network-Cloud services constructed via service composition, which must be realized based on network service description and discovery. In the rest of this article, we will give a survey on state of the art of network service description, discovery, and composition technologies, and discuss the challenges brought in by network-Cloud convergence to these technologies and opportunities for future research in these areas. The focus of our survey is not on Web or Cloud services in general but particularly on NaaS toward network-Cloud convergence.

VI. Network Service Description

Network service description forms the basis of NaaS for network-Cloud convergence as it determines the information that a network service needs to expose for enabling its unambiguous identification and usage in a Cloud environment. For example, in the use case that an application utilizes the Cloud for storing and processing data, network service providers such as AT&T, Verizon, and Comcast all can publish descriptions to provide information about their network services such as provider of the service, source/destination nodes that can be reached by the network, available bandwidth on network routes, achievable delay performance, network access policies, etc.

A. State of the Art

A number of description languages have been proposed for Web service description, some of them have reached the status of standard while others are still the subject of research. W3C and OASIS are the two leading standardization bodies in this area. Web Service Description Language (WSDL), now in its 2.0 version [56], served as the first standard for describing Web services. However, WSDL focuses on syntactic description of Web service interfaces and lacks the ability to provide semantic information. Syntactic descriptions either make service semantics ambiguous or requires a syntax-level agreement between service providers and users, which is too restrictive to fully support the loose-coupling feature required by SOA.

Various technologies have been developed for describing semantic service information. Service semantics are made explicit by reference to a structured vocabulary of terms (ontology) representing a specific area of knowledge. Web Ontology Language (OWL) [57] is a W3C standard for formal ontology description. SAWSDL [58] is the W3C recommendation for adding semantic annotations to WSDL and XML schema. Non-functional service properties, such as reliability, performance, and security that are termed as QoS in general, are also an important aspect of service description. Development progress has been made in producing languages and models for general or domain-specific QoS description. Web Service Quality Model (WSQM) [59] is an ongoing standardization effort of OASIS for Web service QoS specification. Web Service Quality Description Language (WS-QDL) is a description language recommended by W3C for representing QoS by applying the model specified in WSQM.

Various approaches have been proposed for service description of RESTful Web services, but none of them has gained broad support so far. Among the proposals, Web Application Description Language (WADL) [60] seems to be the most mature one. WADL is closely related to WSDL by generating a monolithic file containing all the information about the service interface. In addition, HTML for RESTful Services (hRESTS), which enriches a human-readable document with microformats to make it machine readable, also offers a promising approach to describing RESTful services [61].

The above service description technologies are mainly for general Web services without particularly considering network-as-a-service. Efforts for applying Web service description to networking systems have been made by the telecom community. Open Service Access (OSA) Parlay X [14] and Open Mobile Alliance (OMA) Service Environment (OSE) [19] are two standards for exposing telecom infrastructure to upper layer applications through a service abstraction layer. European Computer Manufacturers Association (ECMA) also published a service description of computer supported telecommunication applications [62]. These standards all employ WSDL for describing network functionalities as Web services, which facilitates creation and deployment of Web services-based telecom applications. However, adoption of basic WSDL in these standards limits their capabilities of describing rich semantic and QoS information of network services. In addition, these specifications mainly focus on telephony-related functions and therefore need to be further developed in order to support the wide variety of multimedia network services in a Cloud environment.

Recent research tends to apply available semantic Web techniques in the network realm and develop ontology specifically for describing network services in a machine-readable format. Network Description Language (NDL) reported in [63] is an ontology based on Resource Description Framework (RDF) [64] designed to describe network elements and topologies. NDL aims to describe an overview of network topology in order to provide a common semantic to the applications, the network, and the service providers for unambiguous communications among them. NDL has been adopted by OGF

NML-WG (Network Mark-up Language Working Group) as the basis for a specification of network topology description [65]. However, the current usage of both the NDL language and the OGF specification focuses on connection-oriented optical networks although extension to the general networking required by Cloud computing is possible.

Network Resource Description Language (NRDL) is developed in [66] in order to facilitate abstraction of networking resources. NRDL is also based on RDF but with a focus on interaction among network elements rather than individual network objects like devices, links, data flow, etc. Such a shift of emphasis enables NRDL to give a better description of the resources for network service provisioning; thus expressing network service abstraction. Resource abstraction of NRDL allows it to be used as a network service description and information exchange protocol between various networking systems and applications.

In addition to topology and connectivity information, networking capabilities and QoS properties are important aspects of network service descriptions. An algorithm for network resource abstraction was given in [67] for describing both network topology and QoS information. This algorithm first abstracts network topology as a full-mesh representation that consists only of service end nodes, then associates the network connectivity between each pair of end nodes with QoS metrics such as bandwidth, delay, and jitter. A capability matrix was developed in [68] for modeling network service capabilities using a capability profile for each ingress-egress pair supported by the service. The profile is defined based on the service curve concept in network calculus theory; thus offering a general service capability description that is independent of network implementations.

Support of virtualization is very important to network service description for NaaS-based network-Cloud convergence. Virtual eXecution Infrastructure Description Language (VXDL) [69] is a language developed for network and computing resource description with virtualization support. This language focuses on describing virtual infrastructure that consists of computing resources interconnected by communication networks. Several features that are important to network virtualization are not sufficiently supported in VXDL schema, for instance no separate entity for network infrastructure description required by InPs to advertise infrastructure resource information to SPs. In order to provide better support for service provisioning in network virtualization environments, the resource description schema proposed in [49] includes entities of substrate resources and virtual resources, and adds attributes that are necessary for supporting a network virtualization framework for IP infrastructure provisioning. A virtual resource description schema was also presented in [70], which integrates the specification of virtual networking resources properties and relations into network service description.

A key aspect of network-Cloud convergence is coordinated control of both networking and computing resources, which requires a service description language that is applicable to heterogeneous types of resources. Toward this end the authors of [71] developed a description language, NDL-OWL, by extending NDL with a more powerful ontology defined in OWL. Using the extensible feature of the RDF/OWL approach, the authors added new types of objects and relationships into the NDL dictionary for describing different types of resources. Currently NDL-OWL can describe various computing capabilities as well as networking resources.

B. Challenges and Research Opportunities

NaaS-based network and Cloud convergence brings in new challenges to service description with respect to the amount and variety of information that needs to be exposed by network services in a Cloud environment. In spite of the aforementioned research progress, this area is still on an early stage and offers numerous research opportunities.

The large scale networking systems for Cloud service provisioning require a balance between richness of information for accurate service description and abstraction of service information for scalable networking. This challenge calls for advances in expressiveness and processing efficiency of service description languages, as well as effective models and algorithms for network resource abstraction and aggregation. Creative combination of the available technologies for network topology aggregation with the latest developments in semantic Web and ontology would be an interesting research direction.

Describing QoS attributes of network services is still an unsolved challenging problem. In contrast to functional service features, there is less agreement regarding the definition and specification of network QoS attributes. Due to the autonomous network domains in a Cloud environment, it is not likely to have a universal approach adopted by all network service providers for describing QoS properties. Measuring QoS parameters of network services is also a difficult task, particularly in the dynamic large scale networks for public Cloud service provisioning. All these factors make QoS description for network services a challenging problem that deserves a thorough investigation.

Cloud computing is expected to employ a wide variety of heterogeneous networking systems. NaaS allows network services to be integrated into a Cloud environment and composed with computing services for Cloud service provisioning. Therefore, research on general description approaches that are applicable to not only heterogeneous network services but also different types of services, including network, Web, and Cloud services, becomes an important and challenging topic.

Since virtualization plays a key role in both Cloud computing and networking and forms a foundation of network-Cloud convergence, network service description language for network-Cloud convergence must have sufficient support for resource virtualization.

The elastic feature of on-demand Cloud service provisioning also brings in a new challenge to network service description. A service description language is used not only by service providers for advertising their service offers but also by service users to specify their service requests. Therefore, NaaS for Cloud computing requires network service description to be able to support dynamic adaptive service specification, which is still an opening topic for future research.

VII. Network Service Discovery

Service discovery plays a key role in NaaS-based network-Cloud convergence by discovering and selecting the network services that meet the requirements for Cloud service provisioning. For example, in the use case that an application utilizes the Cloud for storing and processing data, the application submits a service request that specifies the computing capacity needed for data processing, disk space required for data storing, and the minimum throughput required for data transmission. Then the broker will discover the available Cloud and network services and select Amazon EC2 that meets the computing requirement, Amazon S3 for data storage, a network service provided by Comcast that meets the throughput requirement between the application host and S3 disk, and a network service offered by Verizon for supporting communications between EC2 and S3 servers.

A. State of the Art

Early specifications for service discovery in networking environments include IETF Service Location Protocol (SLP) and industry standards such as Jini, UPnP, Salutation, and Bluetooth. A survey and taxonomy of service discovery protocols in pervasive computing environments is presented in [72]. These protocols are mainly designed for personal/local or enterprise computing environments thus may not scale well to the Internet-based Cloud environment. Service discovery is also an integral part in peer-to-peer (P2P) networks. Various technologies have been developed for scalable and reliable service discovery in large scale P2P networks. An overview of existing solutions for service and resource discovery for a wide variety of networks is given in [73], which includes a detailed discussion on service discovery in peer-to-peer overlay networks. Mobile ad hoc networks bring in special challenges to service discovery and trigger extensive study. [74] gives a survey of research on service advertising, discovery, and selection for mobile ad hoc networks. An overview and comparison of service discovery protocols for multihop mobile ad hoc networks can be found in [75].

Most of the aforementioned technologies focus on locating devices that host functions or contents (e.g., data/files) in a networking environment; which is essentially discovery of computing services in networks rather than discovery of networking resources/capabilities as services. Therefore, these technologies are not directly applicable to NaaS that provides virtualization of networking resources. Nevertheless the obtained results in these areas provide insights on various aspects of service discovery that are valuable to network service discovery in NaaS.

Service discovery has attracted extensive study in the area of Web services. The baseline approach to Web service discovery is the OASIS standard Universal Description, Discovery, and Integration (UDDI) [76], which specifies a data-model for organizing service information with APIs for publishing and querying service descriptions. UDDI has been applied in telecom systems, for example Parlay X and OMA OSE specifications adopted UDDI as the technology for publishing and discovering network services. Although UDDI served as the de-facto standard for service registry during a certain period of time, its data model and query mechanism lack support for semantic and QoS information; thus significantly limiting its application in network service discovery.

A vast number of diverse Web service discovery approaches have been developed. Most of them are based on UDDI but with extensions in two dimensions: i) enhanced data-model of service information to support semantical or hybrid (combined semantical and syntactic) service discovery; and ii) decentralized registry structures and search protocols to achieve scalable service discovery in complex networking environments. The increasing number of available services with dynamic changes and the complexity of such services require a higher degree of automation for service discovery. A survey on a wide variety of Web service discovery technologies is presented in [77], which also evaluates the existing approaches using a list of criteria of autonomic discovery for service discovery automation.

In principle, UDDI-based mechanisms could also be applied for automatic discovery of RESTful Web services. However, to our best knowledge no development of appropriate platform for automatic RESTful service discovery has been reported in literature. A REST service registry is proposed in [78] but no detail about how services are published and discovered at this registry is given. Currently, typical practice of clients for discovering RESTful services is based on offline approaches, which lack the capability of automatic and dynamic service discovery required by NaaS.

Although significant progress has been made in Web service discovery, NaaS-based network-Cloud convergence requires further development on network service discovery. Due to the wide variety of services involved in a converged networking and Cloud computing environment, network service discovery for realizing the NaaS paradigm must be able to cope with the heterogeneity in services. Heterogeneous service descriptions may be published by different service providers, and consequently service registries should be able to manage them. In addition, diverse discovery protocols could be applied in different service domains; thus requiring a mechanism that enables them to cooperate.

A possible approach to addressing heterogeneity in service information is to develop a general purpose model that can be used for mapping different service descriptions. In [79] the authors proposed the PerSeSyn Service Description Model (PSDM) for semantic mapping between different services and developed the PerSeSyn Service Description Language (PSDL) as a common representation for service descriptions and requests. In order to enable service discovery across network domains with different registry structures and search protocols, researchers have developed approaches based on the idea of providing a middleware layer for mapping and forwarding service queries. The Open Service Discovery Architecture (OSDA) designed in [80] can serve as a middleware for inter-domain service discovery that is independent of domain-specific discovery technologies. PYRAMID-S architecture developed in [81] uses a hybrid peer-to-peer topology to organize service registries and provides a scalable framework for unified service publication and discovery over heterogeneous network domains.

Convergence of networking and Cloud computing calls for a holistic vision of service discovery across autonomous domains, including telecommunications, Internet, Web, and Cloud computing. As a possible approach toward such holistic service delivery, the authors of [82] proposed a prototype architecture for an autonomic service management framework that combines SOA, autonomic computing, and application-oriented networking. This framework employs a peer-to-peer overlay broker network and a distributed service registry for enabling scalable service discovery across heterogeneous network domains. The concept of marketplace was originally proposed in Grid computing as an economy-based environment for trading computing resources among heterogeneous systems. In [49] the marketplace concept was extended to network service discovery to design a broker system that discovers virtual network infrastructure services for meeting Cloud service requirements through the marketplace as a common platform. A game-theoretic method for fair resource allocation for Cloud computing services was reported in [83]. A semantic enhanced service exposure framework was proposed in [84] to support discovery of heterogeneous services, including telecom, Web, device, and user-generated services, in a converged service environment.

Dynamic and adaptive network service discovery for NaaS is important to support extensibility and elasticity of Cloud service provisioning. In [85] technology was developed for dynamically adapting Cloud service requests and provisioning according to user context and resource availability. A utility adaptive cloud service broker mechanism was proposed in [86] with its key functions and algorithms. An evolutionary game theoretic mechanism was developed in [87] for adaptive and stable application deployment in Cloud environments. These methods focus on Cloud services in general rather than NaaS in particular. The authors of [88] investigated the problem of adaptive virtual network provisioning and developed an algorithm for adaptive infrastructure resource allocation to support virtual networks. However its application in NaaS for supporting Cloud computing still needs further study.

Dynamic and adaptive network service discovery relies on the latest service information to find the most appropriate services. However, keeping the latest network service descriptions precise becomes particularly challenging due to the dynamic nature of network states and capacities. In order to address this issue, an event-based subscription-notification mechanism was proposed in [89], which maintains the latest network service information at the registry without causing overwhelming communication overhead between service providers and the registry. Another approach is to split static and dynamic service information, publish only static information at the service registry, and then check the dynamic service states only for discovered candidate services in order to make a decision on service selection [90].

With the rapid development of mobile communications and computing, mobile Cloud computing is becoming a significant extension to the notion of Cloud computing [91]. In the basic service model for mobile Cloud computing, mobile devices host service users (typically mobile applications) that access the Cloud infrastructure through wireless mobile networks. Recent development in mobile Cloud computing enables some additional service models such as mobile-as-a-service provider or mobile-as-a service broker. An example application of the new service models is the Hyrax mobile Cloud [92] developed at Carnegie Mellon University for Cloud computing on Android smart-phones. Hyrax Cloud supports client applications that utilize data and execute computing jobs on networks of smart-phones and heterogeneous networks of phones/servers, which allows service providers and/or brokers to be hosted on mobile devices. Results of the Hyrax project show that processing mobile data in-place and transferring data directly between smart-phones could be more efficient and less susceptible to network limitations than offloading data and processing to remote servers. MAPCloud is a hybrid 2-tier Cloud architecture that consists of both local and public clouds infrastructure for improving performance and scalability of mobile applications [93].

For all these possible Cloud service models, discovery of the appropriate wireless mobile networks for supporting mobile Cloud services is an important issue. Access network discovery has been studied in the wireless networking area. Recent progress includes the 3GPP specification for Access Network Discovery and Selection Function (ANDSF) [94] and the network discovery/selection mechanism defined in the IEEE 802.21 protocol [95]. 3GPP has also started investigating applications of SOA in access network discovery and selection. However currently existing approaches focus on network discovery and selection based on local access capabilities and performance. NaaS for supporting mobile Cloud computing requires extensions of these technologies from local access performance-based network discovery to end-to-end service performance-based network discovery. A research effort toward service-oriented network selection in the future mobile Internet is reported in [96]. Applying this technique to support mobile Cloud services would be an interesting research topic.

B. Challenges and Research Opportunities

Although various research developments have been made on discovery of Web, Cloud, and network services, network service discovery for realizing NaaS to support network-Cloud convergence is still an open area with challenges and research opportunities.

Cloud computing may utilize a wide variety of networking systems that scale from LANs to the Internet. Therefore, developing network service discovery mechanisms that meet the scalability requirement is very important. Though research efforts have been made toward this direction, network evolution to the future diversified Internet that allows virtual networks to coexist and operate in parallel brings new challenges to scalable network service discovery that must be fully addressed.

In a converged networking and Cloud computing environment, services are offered at different levels, providing not only data or business logic but also infrastructure resources and capabilities. A wider range of services, including virtual networks as well as virtual machines/servers, must be supported by the discovery mechanism. The holistic vision of networking and computing resources, which forms the foundation of network-Cloud convergence, requires coordinated service management in general and heterogeneous service discovery in particular, across the networking and computing domains. Therefore heterogeneity in service discovery becomes an important and challenging issue that is worth a thorough study.

QoS-based network service discovery is particularly important for NaaS in order to meet Cloud service requirements. Multiple factors make QoS-aware network service discovery a challenging problem, including the lack of a standard approach to specifying QoS attributes, difficulty in measuring network QoS performance, and diversity of Cloud service requirements on network QoS. In addition, discovery of networking and computing services should be combined in a Cloud environment in order to meet user's requirements on Cloud service provisioning. Therefore, QoS-aware discovery and selection of composite network-Cloud services would be an important research topic.

Special features of Cloud computing, especially elastic on-demand dynamic service provisioning, bring in new challenges to NaaS in a Cloud environment that require dynamic and adaptive network service discovery. Therefore, network service discovery needs to be adaptive at run time to the changes in user requirements as well as dynamic network states and QoS capabilities. Awareness of both user demand and service status is a key to enabling adaptive service discovery, which in turn requires the up-to-date service information available at the registry in order to discover and select the appropriate network services. However due to the large scale networking for Cloud computing, frequent update of network service information will generate a large amount of communication and processing overheads. Therefore, research on effective and efficient methods for updating service information and adapting service requests is crucial to achieve high-performance adaptive network service discovery for Cloud computing.

Mobility issues in a converged networking and Cloud computing environment also cause new challenges to network service discovery; thus offering an interesting research topic. Two dimensions of mobility (both physical and logical) must be considered for network-Cloud convergence. With the rapid integration of wireless mobile networks in Cloud computing, mobile devices can be used for accessing as well as hosting Cloud services. This is the physical mobility that network service discovery must be able to cope with. In some special mobile Cloud environments, for example the Hyrax Cloud constructed upon networks of smart-phones, service consumers, service providers, and service brokers may all be hosted on mobile devices, which makes the service discovery problem even more complex. The other dimension of mobility in a Cloud environment is caused by virtual machine/servers migration, which can be referred to as logical mobility. With the convergence of networking and Cloud computing, virtual machines can easily migrate not only within a Cloud data center but also across different data centers through the networks in between, which makes service discovery in such an environment a challenging problem that deserves a thorough investigation.

VIII. Network Service Composition

Cloud service provisioning consists of data communications offered by network infrastructure as well as computing functions provided by Cloud infrastructure. The services offered to end users in a Cloud environment are essentially composite network-Cloud services. Therefore, service composition plays a key role in eventually realizing the NaaS-based network-Cloud convergence. For example, in the use case where an application utilizes the Cloud for storing and processing data, the Cloud services (Amazon EC2 for computing and Amazon S3 for data storing) and network services (Comcast network for data transmission between the application host and S3 disk and Verizon network for communications between S3 and EC2 servers) are composed into a composite end-to-end service that meets the application requirements.

A. State of the Art

Service composition has been an active research area for years in the field of Web services. Numerous technologies have been developed to achieve functional and/or performance requirements of Web service composition. Most of the technologies are based on either workflow management or AI-planning approach, and employ heuristic search, linear programming, or automatic reasoning algorithms. Surveys of recent research on Web service composition can be found in [97] and [98]. A number of conceptual models and corresponding languages have also been developed for modeling and describing Web service composition at different levels. Web Service business Process Execution Language (WS-BPEL) [21] is an OASIS standard that has been widely accepted by industry for modeling Web service composition. Web Service Choreography Description Language (WS-CDL) [99] is a W3C candidate recommendation for Web service composition specification.

For RESTful Web services, elementary building blocks can be composed into more complex services through mashups, which allow data, presentations, or functionalities from two or more resources to be combined for creating new services [100]. However, without a formal service description standard such as WSDL, RESTful services cannot be composed automatically. The lack of an automatic service discovery mechanism also limits the number of RESTful services that can be composed. In addition, mashup-based service composition is mainly static and lacks support of dynamic service composition.

There is a line of research that focuses on Web service composition with respect to QoS requirements, which is particularly important to network services. QoS-aware composition has been studied in the Web services area under different assumptions. If a process template is pre-defined, the objective is to select a service for each task in the process so that the resulting process can meet the given QoS requirements. Optimization of multiple QoS criteria in service composition can be modeled as a multi-dimension multi-choice 0-1 knapsack problem and solved by integer programming [101] or by heuristic search methods [102]. Genetic Algorithms (GA) offer another approach to addressing this problem with the advantage of being able to handle nonlinear QoS constraints [103]. Another type of QoS-aware composition problem is to determine a business process that optimize the end-to-end service QoS criteria. In [104] the authors model both functional and non-functional requirements into integer constraints and propose an integer programming technique to solve the problem. A distributed workload management system was developed in [105] to achieve Web service differentiation and overload protection for QoS provisioning. A recent survey on QoS-aware Web service composition can be found in [106].

The aforementioned research results are developed mainly with Web services as the target. Network service composition is a relatively new concept in the networking domain but has started attracting attention of the research community in this field. In the Ambient Network (AN) project, the notion of dynamic network composition was proposed in order to enable on-demand and transparent cooperation between heterogeneous networks [107]. However, realization of network composition developed in the AN project was based on Generic Ambient Signaling that lacks support of the SOA principle, which limits its application in the NaaS paradigm for supporting Cloud computing [108].

Web service composition technologies have been adopted in the networking domain for network service composition. OMA OSE supports service composition through a mechanism defined as Policy Enforcer. Service enablers in OSE can be implemented in form of Web services as specified in OSE Web Service Enabler Release (OWSER) [109]. Though service composition is not explicitly standardized in OWSER, BPEL [21] is adopted in OSE as a technology to express policy for service orchestration. This opens a door to apply Web service composition technologies for composing network services in OSE; thus allowing network services to be composed through the same mechanism as Web/Cloud services in a Cloud environment. However, the BPEL-based approach is mainly for static service composition and lacks sufficient support of dynamic network service composition required by Cloud computing.

A business model for dynamic composition of web-based telecom services was developed in [110]. The architecture and technologies for realizing this model for dynamic service composition were reported in [111]. In [112], the authors proposed an adoption of loosely coupled service composition into the NGSON framework to integrate existing communications and data services. A general framework for dynamic composition of network services for end-to-end information transport was proposed in [113]. A comparative analysis between SOA and the Open Systems Interconnection (OSI) was presented in this paper to show that the service-orientation principle has existed in the network layering model for years. However, services in OSI are defined and invoked statically thus cannot not be recombined. There are some research efforts toward dynamic composition of network protocol stacks and layers in order to break the limitation of OSI for cross-layer interaction. The idea behind research in this direction can be seen as an attempt to apply the SOA principle in network architecture and employ service composition to enable more flexible interaction and collaboration among network modules.

Network Functional Composition (NFC) is a “clean slate” approach to constructing flexible future Internet architecture. Technologies for realizing NFC have been developed in various projects. These projects include the EU sponsored Autonomic Network Architecture (ANA) project that realized the first non-layered network architecture, the Net-Silo project funded by NSF with an objective to separate control functions from data transport to enable cross-layer interaction, and the 4WARD project for investigating network virtualization to enable concurrent virtual networks on shared infrastructure. The technical approaches of these projects share a common idea of decomposing the layered network stack to functional blocks and organizing these blocks in a composition framework [114]. Through NFC, network services can be composed based on specific application requirements; thus enabling customized networking modules to be assembled into different network services for supporting Cloud computing. However, due to its disruptive approach, thorough investigations on the effectiveness and performance of NFC are still needed before it can be widely adopted into the Internet architecture and Cloud computing environments.

With progress on service compositions in the areas of Web services and networking, composition across these two types of services would be naturally the next step toward a network-Cloud convergence. A comparative study on telecom service composition and Web service composition was conducted in [115], which shows that compositions of telecom services and Web services are not irreconcilable although significant difference exists in these two domains. The authors also examined a number of existing technologies that might be able to support converged telecom-Web services, including Parlay X API, Web service Initiation Protocol (WIP) [116], and ECharts for SIP servlets [117], but found that they only build individual converged applications rather than offering a unified composition framework for converged services. Ericsson Composition Engine reported in [118] is a system for composing converged services crossing the Web, enterprise, and telecom domains, which currently supports Web services as well as SIP in an MIS context and CAP/INAP in an IN context. However SIP may not be widely adopted in the future Internet due to its weak inter-domain capability, which limits the ability of this composition engine for supporting converged Internet-Cloud services. Also, performance of this engine for composing virtual network infrastructure services and Cloud services is not reported in literature yet.

Network-Cloud convergence also demands a language that is able to efficiently specify service composition in heterogeneous service environments, which consist of computing as well as networking services. The existing composition languages, such as WS-BPEL and WS-CDL, are developed mainly for workflow centric and stateless Web services. However, telecom/Internet services are highly event-driven and often require stateful implementations. In order to address this issue, a new language called SCALE [119] has been developed by Ericsson with the goal of supporting heterogeneous service composition. The SCALE language is based mainly on the requirements from the telecom domain, which makes its composition method significantly different than those of languages that are well established and standardized (like WS-BPEL). The general requirements for such a heterogeneous composition language were examined in [120] and then an outline and characteristics of a composition language for heterogeneous service landscape were provided. This proposed language extends BPEL by enhancing its capabilities for supporting event-driven/stateful services in order to enable inter-domain service composition.

In Next Generation Service Overlay Network (NGSON) that was recently standardized by IEEE, function entities for service composition and routing are defined to establish service routes that combine composition of computing services and routing for data communications. The NGSON standard provides a protocol framework for these functional entities without specifying any technology for realizing the protocols. The problem of combined service composition and network routing was studied in [121] with a goal to find the optimal service composition in a network that leads to the minimum routing cost. A decision making system was developed in the paper to solve this problem with AI-planning techniques. Although the authors studied this problem in the context of general networking without considering Cloud computing in particular, the proposed technique may be applied to service composition for network-Cloud convergence.

The elastic on-demand feature of Cloud service provisioning calls for adaptive service composition, which can dynamically adjust the composition according to changes in either service status or user request. Numerous technologies have been developed to address various aspects of adaptive Web services composition. Service substitution, which can be defined as 1-to-1 mapping, was applied in [122] to replace missing services in compositions. [123] proposed the use of a semantic transformer between services to handle semantic mismatching. Conservational service replacement has been tackled by using model checking and graph-based algorithms [124], [125], [126]. People also explored generating mediators (adaptors) between the original composition and new incompatible services, which try to keep the original business process while using new service components [127], [128]. Composition repair and reuse were proposed for fault service diagnosis and recovery. Substitution and compensation actions were developed in [122] to stop process execution when an erroneous condition is encountered and then insert a repair plan by automated reasoning. A language to define substitution actions was proposed in [129]. Dynamically handling service failures was also studied in workflow management [130]. Considering new opportunities motivates study on service composition evolution. The technology developed in [131] alternates composition planning and execution in order to adapt to new opportunities. Plan reuse, which was studied in the AI planning domain, also offers a possible technique for adaptive service composition [132], [133].

Adaptive network service composition has just started attracting research attention in the networking realm and little thorough study on this topic has been reported in the literature. NGSON aims at supporting context-aware dynamic adaptive composition of network services. Although functional entities and protocol mechanism are specified in the NGSON standard, effective implementation technologies are to be fully developed. Dynamic adaptive composition of heterogeneous services (crossing networking and computing domains) is basically still an unexplored area. Investigating possible applications of the existing adaptive Web service composition techniques in the networking field toward enabling adaptive composition of converged network-Cloud services would be an interesting research direction.

B. Challenges and Research Opportunities

Service composition plays a central role in coordinating networking and computing resources to enable the service-oriented network-Cloud convergence. In spite of the extensive study on Web service composition and encouraging research progress for network service composition, service composition for NaaS in a converged networking and Cloud computing environment is still an open problem facing challenges in various aspects; thus offering various interesting topics for future research.

Convergence between the future Internet and Cloud computing naturally leads to an ultra large scale integrated networking and computing environment. Therefore, scalability becomes an important requirement for service composition mechanisms designed for such scenarios, due to the potentially large number of services (both network and Cloud services) involved in compositions. Loose-coupling among services, a key feature of SOA, plays a significant role in scalable service compositions by lowering dependence between service components. Although extensively investigated, designing loosely coupled compositions consisting of an ultra large number of participating services is still an open challenge. It is not yet clear how to combine the needs to compose multiple services, maintain end-to-end QoS, and keep the whole composition coupling level as low as possible.

Heterogeneity is a particularly challenging issue to network service composition for network-Cloud convergence. Networking in a Cloud environment typically comprises heterogeneous network services provided by different autonomous Internet domains. In addition, Cloud services offered to end users comprise services provided by both Cloud infrastructure for computing/storage and network infrastructure for data communications. Therefore, a key aspect of NaaS for Cloud computing is service composition between computing and networking services. Thus, service composition must be realized in a manner independent of the implementation details of each involved service component. However, until recently research on service composition has been conducted separately in the areas of Web services and networking, each of which has its own features and requirements that lead to different technical approaches. Network and Cloud convergence through the NaaS paradigm calls for bridging these two areas toward a general composition framework in which different types of services, including both networking and computing services, are orchestrated to meet Cloud service requirements. This becomes an important and challenging issue for future research.

QoS-aware network service composition is very important for Cloud service provisioning. This becomes particularly challenging for compositions across not only heterogeneous network services but also between computing and networking services. Inter-domain QoS, which has been a difficult problem in networking research for years, becomes more complex in a converged networking and Cloud computing environment due to the coexistence of different types of services. This requires a general composition approach that is agnostic to service types and implementations; thus is applicable to both networking and computing services. Though significant progress

TABLE I
A SUMMARY OF STATE OF THE ART OF SERVICE DESCRIPTION, DISCOVERY, COMPOSITION OF NAAS FOR NETWORK-CLOUD CONVERGENCE

	focus on a single service domain (Web services / computing services)	support network-as-a-service (NaaS)	applicable to heterogeneous services (composite network--Cloud services)	support Cloud service features (adaptive, elastic, mobile, etc.)
Service Description	WSDL [56] SAWSDL [58] WSQM [59] WS description survey [8]	NDL [63], [65] NRDL [66] Parlay X [14] OSE [19] CSTA [62]	NDL-OWL [71] virtual network description [69], [70] PSDL [79]	dynamic network description [90]
Service Discovery	UDDI [76] WS discovery survey [77] service discovery in networks [72], [73], [74], [75]	PSDM [79] OSDA [80] Pyramid-S [81]	holistic service framework [82] service marketplace [49] service exposure model [84]	dynamic adaptive discovery [85], [86], [88] context-aware discovery [89], [90]
Service Composition	WS-PBEL [21] WS-CDL [99] WS composition surveys [98], [97], [106]	network composition [107], [108] OWSER [109] telecom service composition [110], [111], [113] network function composition [114]	NGSON composition [22], [25] combined service composition and routing [24], [121] composition language for heterogeneous services [119], [120] telecom-Web service composition [115], [118]	adaptive service composition [122], [123], [124], [125], [126]

has been made for QoS-aware Web services composition, little work particularly addresses the problem of QoS-aware composition between the networking and Cloud computing service domains. Evaluation of end-to-end QoS of composite network-Cloud services is an unsolved issue and there lacks techniques for QoS modeling and analysis that are applicable to heterogeneous service systems. In addition, virtualization of network and Cloud infrastructure exposes system resources through abstract interfaces, which adds complexity to composite service QoS analysis. Balance between QoS optimization, functional satisfaction, and system scalability is also an issue that is worth studying.

The elastic on-demand feature of Cloud services requires dynamic adaptive network service composition. In spite of the active research on adaptive composition of Web services, as summarized in the previous subsection, little work has been reported on adaptive composition of network-Cloud services. Applications of the available adaptive Web services composition technologies in NaaS for network-Cloud convergence would be an interesting topic for future research. QoS-aware composition adaptation brings in new challenges. Adaptation to QoS attributes by monitoring and predicting service performance along the time would also be useful for supporting elasticity of Cloud service provisioning. On the other hand, QoS equivalence and improvement could be a constraint when selecting substitute services in composition adaptation. Fast adaptation algorithms also need to be developed in order to allow composite network-Cloud services to be adaptive to real-time changes in network and Cloud infrastructure. Therefore, QoS-aware adaptive network-Cloud service composition is a challenging problem that offers rich research opportunities.

Security is a very important aspect in NaaS for network-Cloud convergence. Key requirements include description of network service security attributes, specification of user security requests, security-aware network service discovery, and network service composition for meeting Cloud service security requirements. It is important and challenging to develop a comprehensive framework for inter-domain security policy integration and secure network-Cloud service composition. Security has been such an active research topic in both networking and Cloud computing that this area itself deserves a dedicated survey. Therefore, we leave a thorough survey on the security issue of NaaS for Cloud computing to a future paper.

A summary of state of the art of network service description, discovery, and composition for NaaS-based network-Cloud convergence is given in Table I. The table indicates that service description, discovery, and composition technologies have been extensively studied in the Web services area and has attracted research interest in the telecom/networking field as well. However, most of the existing technologies were developed with a focus on a single (either computing or networking) service domain. Although encouraging research progress has been made on NaaS technologies that are applicable to heterogeneous service environments, converged network-Cloud service provisioning that meets the requirements of Cloud computing is still an open issue.

IX. CONCLUSION

The significant role that networking plays in Cloud computing calls for a holistic vision that allows combined management, control, and optimization of both networking and computing resources in a Cloud environment, which leads to a convergence of networking and Cloud computing. Network virtualization is being adopted in both telecommunications and the Internet as a key attribute for the next generation networking. Virtualization, as a potential enabler of profound changes in both communications and computing domains, is expected to bridge the gap between these two fields and enables a convergence of networking and Cloud computing. The Service-Oriented Architecture (SOA) offers an effective architectural principle for system integration and has been adopted in Cloud service provisioning. The review presented in this article about recent progress in telecom and Internet service provisioning indicates that SOA has been widely applied as a key mechanism for realizing network virtualization. Therefore, SOA, when applied in both Cloud computing and network virtualization, may greatly facilitate the convergence of networking and Cloud computing through the Network-as-a-Service (NaaS) paradigm. Web service technologies, as the main implementation method for SOA, will form the technical foundation of NaaS for network-Cloud convergence. The survey given in this article on key technologies for NaaS, mainly focusing on network service description, discovery, and composition, shows that although significant developments have been made toward enabling NaaS for Cloud computing, this field is still on an early stage and is facing a lot of challenges; thus offering numerous opportunities for future research. Cross-fertilization among multiple areas including telecommunications, computer networking, Web services, and Cloud computing, may provide innovative solutions to network-Cloud convergence that will significantly enhance the performance of the Cloud-based future information infrastructure.

REFERENCES

[1] I. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud computing and grid computing 360-degress compared," in *Proc. 2008 Grid Computing Environment Workshop*, pp. 1–10.

[2] K. R. Jackson, K. Muriki, S. Canon, S. Cholia, and J. Shalf, "Performance analysis of high performance computing applications on the Amazon Web services Cloud," in *Proc. 2010 IEEE International Conference on Cloud Computing Technology and Science*, pp. 159–168.

[3] G. Wang and T. S. E. Ng, "The impact of virtualization on network performance of Amazon EC2 data center," in *Proc. 2010 IEEE INFOCOM*, pp. 1–9.

[4] T. Magedanz, N. Blum, and S. Dutkowski, "Evolution of SOA concepts in telecommunications," *IEEE Computer Mag.*, vol. 40, no. 11, pp. 46–50, 2007.

[5] N. M. M. K. Chowdhury and R. Boutaba, "Network virtualization: state of the art and research challenges," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 20–26, 2009.

[6] T. Erl, *Service-Oriented Architecture – Concepts, Technology, and Design*. Prentice Hall, 2005.

[7] D. Griffin and D. Pesch, "A survey on Web services in telecommunications," *IEEE Commun. Mag.*, vol. 45, no. 7, pp. 28–35, 2007.

[8] V. Issarny, N. Georgantas, S. Hachem, A. Zarras, P. Vassiliadist, M. Autili, M. A. Gerosa, and A. B. Hamida, "Service-oriented middleware for the future Internet: state of the art and research directions," *J. Internet Services and Applications*, vol. 2, no. 1, pp. 23–45, 2011.

[9] K. Channabasavaiah, K. Holley, and E. Tuggle, "Migrating to a service-oriented architecture," *IMB DeveloperWorks*, Dec. 2003.

[10] OASIS, "Reference model for service-oriented architecture 1.0. Available: http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf," Oct. 2006.

[11] S. Weerawarana, F. Curbera, F. Leymann, T. Storey, and D. F. Ferguson, *Web Services Platform Architecture*. Prentice Hall, 2005.

[12] M. Lanthaler and C. Gutl, "Toward a RESTful service ecosystem," in *Proc. 2010 IEEE International Conference on Digital Ecosystems and Technologies*, pp. 209–214.

[13] T. Magedanz, "IN and TMN: providing the basis for future information networking architectures," *Computer Commun. J.*, vol. 16, no. 5, pp. 267–276, 1993.

[14] ETSI, "Parlay X 3.0 Specifications. Available: http://docbox.etsi.org/tispan/open/osa/parlayx30.html," Nov. 2007.

[15] ITU-T, "Rec Y.2012: Functional Requirements and Architecture of the NGN release 1," Sep. 2006.

[16] 3GPP, "IP Multimedia Subsystem (IMS), stage 2, Technical Specification 23.228," 2006.

[17] K. Knightson, N. Morita, and T. Towl, "NGN architecture: generic principles, functional architecture, and implementation," *IEEE Commun. Mag.*, vol. 43, no. 10, pp. 49–56, 2005.

[18] ITU-T, "Rec Y.2240: Requirements and Capabilities for NGN Service Integration and Delivery Envrionment," May 2006.

[19] OMA, "Open Service Environment version 1.0. Available: http://www.openmobilealliance.org/technical," Oct. 2009.

[20] TMForum, "TMF061 Service Delivery Framework (SDf) reference architecture," July 2009.

[21] OASIS, "Web Services Business Process Execution Language (WS-BPEL) version 2.0. Available: http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.pdf," Apr. 2007.

[22] IEEE, "Standard 1903: Functional Architecture of the Next Generation Service Overlay Networks," Oct. 2011.

[23] C. Makaya, A. Dutta, B. Falchuk, D. Chee, S. Das, F. Lin, M. Ito, S. Komorita, T. Chiba, and H. Yokota, "Enhanced next-generation service overlay networks architecture," in *Proc. 2010 IEEE International Conference on Internet Multimedia Systems Architecture and Application*, pp. 1–6.

[24] C. Makaya, B. Falchuk, D. Chee, F. Lin, S. Das, M. Ito, S. Komorita, T. Chiba, and H. Yokota, "Service composition based on next-generation service overlay network architecture," in *Proc. 2011 IFIP International Conference on New Technologies, Mobility, and Security*, pp. 1–6.

[25] S. lk Lee and S.-G. Kang, "NGSON: features, state of the art, and realization," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 54–61, 2012.

[26] T. Anderson, L. Peterson, S. Shenker, and J. Turner, "Overcoming the Internet impasses through virtualization," *IEEE Computer Mag.*, vol. 38, no. 4, pp. 34–41, 2005.

[27] GENI-Planning-Group, "GENI design principles," *IEEE Computer Mag.*, vol. 39, no. 9, pp. 102–105, 2006.

[28] J. Turner and D. E. Taylor, "Diversifying the Internet," in *Proc. 2005 IEEE Global Communications Conference*, pp. 755–760.

[29] N. Feamster, L. Gao, and J. Rexford, "How to lease the Internet in your spare time," *ACM SIGCOMM Computer Commun. Rev.*, vol. 37, no. 1, pp. 61–64, 2007.

[30] N. Niebert, S. Baucke, I. El-Khayat, M. Johnsson, B. Ohlman, H. Abramowica, K. Wuenstel, H. Woesner, J. Ouittek, and L. M. Correia, "The way 4WARD to the creation of a future Internet," in *Proc. 2008 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5.

[31] P. Szegedi, J. F. Riera, J. A. Garcia-Espin, M. Hidell, P. Sjodin, P. Soderman, M. Ruffini, D. O'Mahony, A. Bianco, L. Giraudo, M. P. de Leon, G. Power, C. Cerveddo-Pastor, V. Lopez, and S. Naegele-Jackson, "Enabling future Internet research: the FEDERICA case," *IEEE Commun. Mag.*, vol. 49, no. 7, pp. 54–61, 2011.

[32] M. Boucadair, P. Georgatsos, N. Wang, D. Driffin, G. Pavlou, and A. Elizondo, "The AGAVE approach for network virtualization: differentiated services delivery," *Springer Ann. Telecommun. J.*, vol. 64, no. 5, pp. 277–288, 2009.

[33] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: enabling innovation in campus networks," *ACM SIGCOMM Computer Commun. Rev.*, vol. 38, no. 2, pp. 69–74, 2008.

[34] E. Grasa, G. Junyent, S. Figuerola, A. Lopez, and M. Savoie, "UCLPv2: a network virtualization framework built on Web services," *IEEE Commun. Mag.*, vol. 46, no. 3, pp. 126–134, 2008.

[35] S. Figuerola and M. Lemay, "Infrastructure services for optical networks," *IEEE/OSA J. Optical Commun. and Networks*, vol. 1, no. 2, pp. A247–257, 2009.

[36] G. Branca, P. Anedda, and L. Atzori, "Transport stratum services in NGN: a SOA-oriented design," in *Proc. 2010 IEEE Global Communication Conference*, pp. 1–5.

[37] M. E. Barchi, N. Kara, and R. Dssouli, "Toward a service-oriented network virtualization architecture," in *Proc. 2010 ITU-T Kaleidoscope Conference*, pp. 1–7.

[38] Q. Duan, "End-to-end modelling and performance analysis for network virtualisation in the next generation Internet," *International J. Commun. Networks and Distribut. Syst.*, vol. 8, no. 1, pp. 53–69, 2012.

[39] Q. Duan, "Analysis on quality of service provisioning for communication services in network virtualization," *J. Commun.*, vol. 7, no. 2, pp. 143–154, 2012.

[40] Cisco, "Using infrastructure service orchestration to enable a service-oriented architecture," 2009.

[41] L.-J. Zhang and Q. Zhou, "CCOA: cloud computing open architecture," in *Proc. 2009 Symposium on Network System Design and Implementation*.

[42] W. Tsai, X. Sun, and J. Balasooriya, "Service-oriented cloud computing architecture," in *Proc. 2010 International Conference on Information Technology: New Generations*, pp. 684–689.

[43] Y. Wei and M. B. Blake, "Service-oriented computing and cloud computing: challenges and opportunities," *IEEE Internet Computing*, vol. 14, no. 6, pp. 72–75, 2010.

[44] OGF, "Open Cloud Computing Interface." Available: http://occi-wg.org/, May 2010.

[45] IRMOS, "Intelligent Service Oriented Network Infrastructure." Available: http://www.irmosproject.eu/isoni.aspx, July 2009.

[46] NSF, "FIA Collaborative Research: NEBULA: a future Internet that supports trustworthy cloud computing." Available: www.nsf.gov/awardsearch/showaward.do?awardnumber=1040672, Sep. 2010.

[47] SAIL, "Cloud Network Architecture Description." Available: http://www.sail-project.eu/deliverables/," July 2011.

[48] S. Nepal, J. Chan, S. Chen, D. Moreland, and J. Zic, "An infrastructure virtualization SOA for VNO-based business models," in *Proc. 2007 IEEE International Conference on Services Computing*, pp. 44–51.

[49] B. Peng, A. Hammad, R. Nejabati, S. Azodolmolky, and V. Reijs, "A network virtualization framework for IP infrastructure provisioning," in *Proc. 2011 IEEE International Conference on Cloud Computing Technology and Science*, pp. 679–684.

[50] J. Matias, E. Jacob, D. Sanchez, and Y. Demchenko, "An OpenFlow based network virtualization framework for the cloud," in *Proc. 2011 IEEE International Conference on Cloud Computing Technology and Science*, pp. 672–678.

[51] Y. Zhang, K. Zheng, K. Chen, C. Hu, and A. V. Vasilakos, "CoSwitch: a cooperative edge switch design for software defined data center networking," in *Proc. 2013 IEEE International Conference on Computer Communications*.

[52] C. Shen, C. Heng, and X. Zhou, "Inter-cloud operations via NGSON," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 82–89, 2012.

[53] Cisco, "The dynamic network cloud: provision and orchestrate services with integrated Cisco and BMC solution," 2010.

[54] ITU-T, "Focus Group on Cloud Computing (FG-Cloud)." Available: http://www.itu.int/en/itu-t/focusgroups/cloud/pages/default.aspx, Apr. 2007.

[55] OGF, "OGF-WG working draft: Network Service Framework v1.0," Dec. 2010.

[56] W3C, "Web Service Description Language (WSDL) version 2.0." Available: http://www.w3.org/tr/wsdl20-primer/, June 2007.

[57] W3C, "OWL 2 Web Ontology Language Document Overview," Oct. 2009.

[58] W3C, "Semantic Annotation for WSDL and XML-Schema." Available: http://www.w3.org/tr/sawsdl/, Aug. 2007.

[59] OASIS, "Web Services Quality Model (WSQM)." Available: http://www.oasis-open.org/committees/, Aug. 2004.

[60] W3C, "Web Application Description Language (WSDL) member submission." Available: http://www.w3.org/submission/wadl/, Aug. 2009.

[61] J. Kopecky, K. Gomadam, and T. Vitvar, "hRESTS: an HTML microformat for describing RESTful web services," in *Proc. 2008 IEEE/WIC/ACM International Conference on Web Intelligent and Intelligent Agent Technology*, pp. 619–625.

[62] ECMA, "Services for Computer Supported Telecommunications Applications (CSTA)," 9th edition. Available: http://www.ecma-international.org/publications/standards/ecma-269.htm, Dec. 2011.

[63] J. Ham, P. Grosso, R. Pol, A. Toonk, and C. Laat, "Using the network description language in optical networks," in *Proc. 2007 IFIP/IEEE International Symposium on Integrated Network Management*, pp. 199–205.

[64] W3C, "Resource Description Framework (RDF)," Feb. 2004.

[65] OGF, "GFD-I.165 Network topology descriptions in hybrid networks," Mar. 2010.

[66] A. Campi and F. Callegai, "Network resource description language," in *Proc. 2009 IEEE Global Communication Conference*, pp. 1–6.

[67] C. E. Abosi, R. Nejabati, and D. Simeonidou, "A novel service composition mechanism for the future optical Internet," *J. Optical Commun. and Networking*, vol. 1, no. 2, pp. A106–A120, 2009.

[68] Q. Duan, "Network service description and discovery for high-performance ubiquitous and pervasive Grids," *ACM Trans. Autonomous and Adaptive Systems*, vol. 6, no. 1, 2011.

[69] G. P. Koslovski, P. V.-B. Primet, and A. S. Charao, "VXDL: virtual resources and interconnection networks description language," in *Proc. 2008 International Conference on Networks for Grid Applications*.

[70] E. Renault, W. Louati, I. Houidi, and H. Medhioub, "A framework to describe and search for virtual resource objects," *Springer Lecture Notes in Computer Sciences Future Generation Information Technology*, vol. 6485, pp. 208–219, 2010.

[71] I. Baldine, Y. Xin, A. Mandal, C. Heermann, J. Chase, V. Marupadi, A. Yumerefendi, and D. Irwin, "Networked cloud orchestration: a GENI perspective," in *Proc. 2010 IEEE Workshop on Management of Emerging Networks and Services*, pp. 573–578.

[72] F. Zhu, M. W. Mutka, and L. M. Ni, "Service discovery in pervasive computing environments," *IEEE Pervasive Computing*, vol. 4, no. 4, pp. 81–89, 2005.

[73] E. Meshkova, J. Riihijarvi, M. Petrova, and P. Mahonen, "A survey on resource discovery mechanisms, peer-to-peer and service discovery frameworks," *Computer Networks J.*, vol. 52, no. 11, pp. 2097–2128, 2008.

[74] C. N. Ververidis and G. C. Polyzos, "Service discovery for mobile ad hoc networks: a survey of issues and techniques," *IEEE Commun. Surveys & Tutorials*, vol. 10, no. 3, pp. 30–45, 2008.

[75] A. N. Mian, R. Baldoni, and R. Beraldi, "A survey of service discovery protocols in multihop mobile ad hoc networks," *IEEE Pervasive Computing*, vol. 8, no. 1, pp. 66–74, 2009.

[76] OASIS, "Universal Description, Discovery and Integration (UDDI) version 3.0.2." Available: http://uddi.org/pubs/uddi-v3.htm, Feb. 2005.

[77] M. Rambold, H. Kasinger, F. Lautenbacher, and B. Bauer, "Toward automatic service discovery–a survey and comparison," in *Proc. 2009 IEEE International Conference on Services Computing*, pp. 192–201.

[78] J. Guo, B. Cheng, J. Chen, and X. Lin, "Applying recommender system based mashup to web-telecom hybrid service creation," in *Proc. 2009 IEEE Global Communications Conference*.

[79] S. B. Mokhtar, P.-G. Raverdy, A. Urbieta, and R. S. Cardoso, "Interoperable semantic & syntactic service matching for ambient computing environments," in *Proc. 2008 International Workshop on Ad-hoc Ambient Computing*.

[80] N. Limam, J. Ziembicki, R. Ahmed, Y. Iraqi, D.-T. Li, R. Boutaba, and F. Cuervo, "OSDA: open service discovery architecture for efficient cross-domain service provisioning," *Computer Commun. J.*, vol. 30, no. 3, pp. 546–563, 2007.

[81] T. Pilioura and A. Tsalgatidou, "Unified publication and discovery of semantic Web services," *ACM Trans. on the Web*, vol. 3, no. 3, 2009.

[82] Y. Cheng, A. Leon-Garcia, and I. Foster, "Toward an automatic service management framework: a holistic vision of SOA, AON, and autonomic computing," *IEEE Commun. Mag.*, vol. 46, no. 5, pp. 138–146, 2008.

[83] G. Wei, A. V. Vasilakos, Y. Zheng, and N. Xiong, "A game-theoretic method of fair resource allocation for cloud computing services," *J. Supercomputing*, vol. 54, no. 2, pp. 252–269, 2010.

[84] C. Huang, G. M. Lee, and N. Crespi, "A semantic enhanced service exposure model for a converged service environment," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 32–40, 2012.

[85] P. Papakos, L. Capra, and D. S. Rosenblum, "VOLARE: context-aware adaptive cloud service discovery for mobile systems," in *Proc. 2010 International Workshop on Adaptive and Reflective Middleware*.

[86] H. Song, C. S. Bae, J. W. Lee, and C.-H. Youn, "Utility adaptive service brokering mechanism for personal cloud services," in *Proc. 2011 Military Communications Conference*, pp. 1622–1627.

[87] C. Lee, J. Suzuki, A. V. Vasilakos, Y. Yamamoto, and K. Oba, "An evolutionary game theoretic approach to adaptive and stabel application deployment in clouds," in *Proc. 2010 Workshop on Bio-Inspired Algorithms for Distributed Systems*.

[88] I. Houidi, W. Louati, D. Zeghlache, P. Papadimitriou, and L. Mathy, "Adaptive virtual network provisioning," in *Proc. 2010 ACM SIGCOMM Workshop on Virtualized Infrastructure Systems and Architectures*, pp. 41–48.

[89] Q. Duan, "Automatic network service discovery and selection in virtualization-based future Internet," in *Proc. 2011 IEEE Global Communication Conference Workshops*, pp. 1088–1093.

[90] U. Kuster and B. Konig-Ries, "Supporting dynamics in service descriptions–the key to automatic service usage," in *Proc. 2007 International Conference on Service Oriented Computing*.

[91] L. Guan, X. Ke, M. Song, and J. Song, "A survey of research on mobile cloud computing," in *Proc. 2011 IEEE/ACIS International Conference on Computer and Information Science*, pp. 387–392.

[92] E. E. Marinelli, "Hyrax: cloud computing on mobile devices using MapReduce," Carnegie Mellon University Technical Report CMU-CS-09-164, Sep. 2009.

[93] M. R. Rahimi, N. Venkatasubramanian, S. Mehrotra, and A. V. Vasilakos, "MAPCloud: mobile applications on an elastic and scalable 2-tier cloud architecture," in *Proc. 2010 IEEE/ACM International Conference on Utility and Cloud Computing*.

[94] 3GPP, "Technical specification Access Network Discovery and Selection Function (ANDSF)," Oct. 2008.

[95] A. D. L. Oliva, A. Banchs, I. Soto, and A. Vidal, "An overview of IEEE 802.21 media-independent handover services," *IEEE Wireless Commun.*, vol. 15, no. 4, pp. 96–103, 2008.

[96] Q. Duan and Y. Lu, "Service-oriented network discovery and selection in virtualization-based mobile Internet," *J. Computer Inf. Systems*, in press.

[97] J. Rao and X. Su, "A survey of automated Web service composition methods," in *Proc. 2004 International Workshop on Semantic Web Services and Web Process Composition*.

[98] S. Dustdar and W. Schreiner, "A survey on Web services composition," *Int'l J. Web and Grid Services*, vol. 1, no. 1, pp. 1–30, 2005.

[99] W3C, "Web service choreography description language version 1.0." Available: http://www.w3.org/tr/ws-cdl-10/, Nov. 2005.
[100] F. Belqasmi, R. Glitho, and C. Fu, "RESTful Web services for service provisioning in next-generation networks: a survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 66–73, 2011.
[101] L. Zeng, B. Benatallah, A. H. H. Ngu, M. Dumas, J. Kalagnanam, and H. Chang, "QoS-aware middleware for Web services composition," *IEEE Trans. Software Engineer.*, vol. 30, no. 5, 2004.
[102] T. Yu, Y. Zhang, and K. Lin, "Efficient algorithms for Web services selection with end-to-end QoS constraints," *ACM Trans. on the Web*, vol. 1, no. 1, 2007.
[103] G. Canfora, M. D. Penta, R. Esposito, and M. L. Villani, "An approach for QoS-aware service composition based on genetic algorithms," *Lecture Notes in Computer Science*, vol. 3387, pp. 43–54, 2005.
[104] L. Cui, S. Kumara, and D. Lee, "Scenario analysis of Web service composition based on multi-criteria mathematical programming," *INFORMS Service Science*, vol. 3, no. 3, 2011.
[105] S. Zhang, H. Wu, W. Wang, B. Yang, P. Liu, and A. V. Vasilakos, "Distributed workload and response time management for Web applications," in *Proc. 2011 International Conference on Network and Services Management*, pp. 1–9.
[106] A. Strunk, "QoS-aware service composition: a survey," in *Proc. 2010 IEEE European Conference on Web Services*, pp. 67–74.
[107] C. Kappler, P. Poyhonen, M. Johnsson, and S. Schmid, "Dynamic network composition for beyond 3G networks: a 3GPP viewpoint," *IEEE Network Mag.*, vol. 21, no. 1, pp. 47–52, 2007.
[108] F. Belqasmi, R. Glitho, and R. Dssouli, "Ambient network composition," *IEEE Network Mag.*, vol. 22, no. 4, pp. 6–12, 2008.
[109] OMA, "Open Mobile Alliance Web Services Enabler version 1.1," Mar. 2006.
[110] R. Karunamurthy, F. Khendek, and R. H. Glitho, "A business model for dynamic composition of telecommunication Web services," *IEEE Commun. Mag.*, vol. 45, no. 7, pp. 36–43, 2007.
[111] R. Karunamurthy, F. Khendek, and R. H. Glitho, "A novel architecture for Web service composition," *Elsevier J. Networks and Computer Applications*, vol. 35, no. 2, pp. 787–802, 2011.
[112] S. Komorita, M. Ito, H. Yokota, C. Makaya, B. Falchuk, D. Chee, and S. Das, "Loosely coupled service composition for deployment of next generation service overlay networks," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 62–72, 2012.
[113] C. Fortuna and M. Mohorcic, "Dynamic composition of service for end-to-end information transport," *IEEE Wireless Commun.*, vol. 16, no. 4, pp. 56–62, 2009.
[114] C. Henke, A. Siddiqui, and R. Khondoker, "Network functional composition: state of the art," in *Proc. 2010 IEEE Australasian Telecommunication Networks and Applications Conference*, pp. 43–48.
[115] G. Bond, E. Cheung, I. Fikouras, and R. Levenshteyn, "Unified telecom and Web services composition: problem definition and future direction," in *Proc. 2009 Conference on Principles, Systems and Applications of IP Telecommunications*.
[116] W. Chou, L. Li, and F. Liu, "Web services for communication over IP," *IEEE Commun. Mag.*, vol. 46, no. 3, pp. 136–143, 2008.
[117] T. M. Smith and G. W. Bond, "ECharts for SIP servlets user manual," *AT&T Techn. Report TD-7BELCD*, 2007.
[118] J. Niemoller, E. Freiter, K. Vandikas, R. Quinet, R. Levenshteyn, and I. Fikouras, "Composition in converged service networks: requirements and solutions," in *Proc. 2010 International Workshop on Business Systems Management and Engineering*.
[119] Ericsson, "Research Report: SCAL–a language for dynamic composition of heterogeneous services," Nov. 2010.
[120] J. Niemoller, K. Vandikas, R. Levenshteyn, D. Schleicher, and F. Leymann, "Towards a service composition language for heterogeneous service environments," in *Proc. 2011 IEEE International Conference on Intelligence in Next Generation Networks*, pp. 121–126.
[121] X. Huang, S. Shanbhag, and T. Wolf, "Automated service composition and routing in networks with data-path services," in *Proc. 2010 IEEE International Conference on Computer Communication Network*, pp. 1–8.
[122] G. Friedrich and V. Ivanchenko, "Model-based repair of web service processes," Tech. Rep. 2008/001, ISBI Research Group, Alpen-Adria-Universität Klagenfurt, 2008.
[123] O. Moser, F. Rosenberg, and S. Dustdar, "Non-intrusive monitoring and service adaptation for WS-BPEL," in *Proc. 2008 International World Wide Web Conference*, pp. 815–824.
[124] L. Cavallaro, E. D. Nitto, and M. Pradella, "An automatic approach to enable replacement of conversational services," in *Proc. 2009 International Conference on Service-Oriented Computing / ServiceWave*, pp. 159–174.
[125] R. Mateescu, P. Poizat, and G. Salaün, "Adaptation of service protocols using process algebra and on-the-fly reduction techniques," in *Proc. 2008 International Conference on Service-Oriented Computing*, pp. 84–99.
[126] D. Grigori, J. C. Corrales, and M. Bouzeghoub, "Behavioral matchmaking for service retrieval: application to conversation protocols," *Inf. Syst. J.*, vol. 33, no. 7-8, pp. 681–698, 2008.
[127] W. M. P. van der Aalst, A. J. Mooij, C. Stahl, and K. Wolf, "Service interaction: patterns, formalization, and analysis," in *Proc. 2009 International School on Formal Methods for the Design of Computer, Communications and Software Systems: Web Services*.
[128] H. R. M. Nezhad, G. Y. Xu, and B. Benatallah, "Protocol-aware matching of web service interfaces for adapter development," in *Proc. 2010 International World Wide Web Conference*, pp. 731–740.
[129] L. Baresi, S. Guinea, and L. Pasquale, "Self-healing BPEL processes with dynamo and the jboss rule engine," in *Proc. 2007 International Workshop on Engineering of Software Services for Pervasive Environments*, pp. 11–20.
[130] M. Gajewski, M. Momotko, H. Meyer, H. Schuschel, and M. Weske, "Dynamic failure recovery of generated workflows," in *Proc. 2005 International Conference on Database and Expert Systems Applications Workshops*, pp. 982–986.
[131] A. Lazovik, M. Aiello, and M. P. Papazoglou, "Planning and monitoring the execution of web service requests," *Int'l J. Digital Libraries*, vol. 6, no. 3, pp. 235–246, 2006.
[132] S. Koenig, D. Furcy, and C. Bauer, "Heuristic search-based replanning," in *Proc. 2002 International Conference on AI Planning and Scheduling*, pp. 294–301.
[133] R. van der Krogt and M. de Weerdt, "Plan repair as an extension of planning," in *Proc. 2005 International Conference on Automated Planning and Scheduling*, pp. 161–170.

Qiang Duan is an Assistant Professor of Information Science and Technology at the Pennsylvania State University Abington College. His current research areas include computer communications and networking, the next generation Internet, network virtualization, service-oriented architecture, Cloud computing, and Web services. He has published more over 50 technical papers in international journals and conference proceedings and authored 5 book chapters. Dr. Duan is serving as an associate editor for Journal of Network Protocols and Algorithms and International Journal of Internet and Distributed Computing Systems. He has served on the technical program committees for numerous conferences including GLOBECOM, ICC, ICCCN, AINA, WCNC, and served as a reviewer for various journals including IEEE JSAC, IEEE TNSM, IEEE TPDS, and ACM TAAS. Dr. Duan received his Ph.D. degree in electrical engineering from the University of Mississippi. He holds a B.S. degree in electrical and computer engineering and a M.S. degree in telecommunications and electronic systems.

Yuhong Yan is an Assistant Professor at the Department of Computer Science and Software Engineering at Concordia University, Montreal, Canada since June 2008. Before joining Concordia, she was a Researcher Officer in the Institute for Information Technology (IIT) in the Canadian National Research Council's (NRC) since Feb. 2003. She is also an Adjunct Professor in the Faculty of Computer Science in the University of New Brunswick, Canada since 2004. Her current research focuses on Service Computing, especially in formal modeling, composition, fault diagnosis, reparation, and adaptation of Web service processes. She has authored over 40 articles and papers. She is one of the organizers of IEEE ICWS and SCC in recent years. Dr. Yan's work on this paper is supported by the Engineering Research Council of Canada (NSERC) Discovery Grant "Service Oriented Systems Integration."

Athanasios V. Vasilakos is currently a Professor at Kuwait University. He has authored or co-authored over 200 technical papers in major international journals and conferences. He is author/coauthor of five books and 20 book chapters in the areas of communications.

Prof. Vasilakos has served as General Chair, Technical Program Committee Chair for many international conferences. He served or is serving as an Editor or/and Guest Editor for many technical journals, such as the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS–PART B: CYBERNETICS, IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, IEEE TRANSACTIONS ON COMPUTERS, *ACM Transactions on Autonomous and Adaptive Systems*, the IEEE JSAC special issues of May 2009, Jan. 2011, and Mar. 2011, the *IEEE Communications Magazine*, *ACM/Springer Wireless Networks (WINET)*, and *ACM/Springer Mobile Networks and Applications (MONET)*. He is founding Editor-in-Chief of the *International Journal of Adaptive and Autonomous Communications Systems (IJAACS)* (http://www.inderscience.com/ijaacs), and the *International Journal of Arts and Technology (IJART)* (http://www.inderscience.com/ijart).

He is General Chair of the Council of Computing of the European Alliances for Innovation.