Computer Networks 93 (2015) 66–79

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Survey Paper

Wireless software-defined networks (W-SDNs) and network function virtualization (NFV) for 5G cellular systems: An overview and qualitative evaluation

Ian F. Akyildiz[a,1], Shih-Chun Lin[a,2,*], Pu Wang[b,3]

[a] *Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA*

[b] *Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS 67260, USA*

ARTICLE INFO

Article history:
Received 11 October 2015
Accepted 17 October 2015
Available online 27 October 2015

Keywords:
Wireless software-defined networks (W-SDNs)
Software-defined networking (SDN)
Network function virtualization (NFV)
Software-defined traffic engineering
5G cellular systems
Cloud radio access network (C-RAN)

ABSTRACT

Cellular network technologies have evolved to support the ever-increasing wireless data traffic, which results from the rapidly-evolving Internet and widely-adopted cloud applications over wireless networks. However, hardware-based designs, which rely on closed and inflexible architectures of current cellular systems, make a typical 10-year cycle for a new generation of wireless networks to be standardized and deployed. To overcome this limitation, the concept of software-defined networking (SDN) has been proposed to efficiently create centralized network abstraction with the provisioning of programmability over the entire network. Moreover, the complementary concept of network function virtualization (NFV) has been further proposed to effectively separate the abstraction of functionalities from the hardware by decoupling the data forwarding plane from the control plane. These two concepts provide cellular networks with the needed flexibility to evolve and adapt according to the ever-changing network context and introduce wireless software-defined networks (W-SDNs) for 5G cellular systems. Thus, there is an urgent need to study the fundamental architectural principles underlying a new generation of software-defined cellular network as well as the enabling technologies that supports and manages such emerging architecture. In this paper, first, the state-of-the-art W-SDNs solutions along with their associated NFV techniques are surveyed. Then, the key differences among these W-SDN solutions as well as their limitations are highlighted. To counter those limitations, SoftAir, a new SDN architecture for 5G cellular systems, is introduced.

1. Introduction

Existing commercial wireless networks are inherently hardware-based and rely on closed and inflexible architectural designs. Such inflexible hardware-based architectures typically lead to a 10-year cycle for a new generation of wireless networks to be standardized and deployed, impose significant challenges into adopting new wireless networking technologies to maximize the network capacity and coverage, and prevent the provision of truly-differentiated services able to adapt to increasingly growing, uneven, and highly variable traffic patterns. In particular, for 5G cellular system requirements, the ultra high capacity should have 1000-fold capacity/km^2 compared to LTE, the user-plane latency should be less than 1ms over the radio access network (RAN), and

* Corresponding author. Tel.:+7654048029.
E-mail addresses: ian@ece.gatech.edu (I.F. Akyildiz), slin88@ece.gatech.edu (S.-C. Lin), pu.wang@wichita.edu (P. Wang).

[1] Fellow, IEEE.
[2] Student Member, IEEE.
[3] Member, IEEE.

http://dx.doi.org/10.1016/j.comnet.2015.10.013
1389-1286/

the ultra high data rates should provide 100-fold increase in user-experienced throughput (targeting 1Gbps experienced user throughput everywhere). The challenges faced by the current network architectures cannot be solved without a radical paradigm shift in the design of next-generation wireless networks. Hence, in this paper, we survey the state-of-the-art solutions that exploit software-defined networking (SDN) in the architecture designs of wireless software-defined networks (W-SDNs) and qualitatively evaluate their performance with a focus on network function virtualization (NFV).

SDN has been recently introduced primarily for data center networks and for the next-generation Internet [1–3]. The main ideas are (i) to separate the data plane from the control plane; and (ii) to introduce novel network control functionalities based on an abstract representation of the network. In current instantiations of this idea, these are realized by (i) removing control decisions from the hardware, e.g., switches, (ii) enabling the hardware to be programmable through an open and standardized interface, e.g., Openflow [4], and (iii) using a network controller to define the behavior and operation of the network forwarding infrastructure. SDN makes it easier to introduce and deploy new applications and services than the classical hardware-dependent standards.

Another important concept in the context of SDN is NFV. As highly complementary to SDN, NFV effectively abstract network functionalities and implement them in software. By such way, network functions, e.g., routing decisions, can be separated from the local devices and be implemented at remote servers or Cloud. SDN and NFV are mutually beneficial but are not dependent on each other. In particular, network function can be virtualized and deployed without a SDN being required and vice versa. NFV brings many potential benefits from cost reduction to great variety of system openness. First, NFV reduces CAPEX, OPEX, and power consumption through consolidating equipment and exploiting the economies of scale of the IT industry. Moreover, it increases speed of time to market by minimizing the typical network operator cycle of innovation. Third, NFV provides the availability of network appliance multi-version and multi-tenancy, which allows uses of a single platform for different applications, users and tenants, which enables a wide variety of eco-systems and encourages openness.

So far, several state-of-the-art solutions already employ these emerging concepts more or less, and introduce integrated architectures of W-SDNs from different design axes for 5G cellular systems. The overview of these solutions as well as their strengths and weakness are summarized and compared with our proposed solution, called SoftAir [3], in details from a qualitative perspective. Regarding our proposed SoftAir architecture, the *control plane* consists of network management and optimization tools and is implemented on the network servers. The *data plane* consists of software-defined base stations (SD-BSs) in the RAN and software-defined switches (SD-switches) in the cellular core network (CN). Their control logic, e.g., physical/MAC/network functions, are implemented in software on general purpose computers and remote data centers. Our proposed SoftAir architecture offers five core properties: (i) *programmability*, i.e., SDN nodes (e.g., SD-BSs and SD-switches) can be reprogrammed on-the-fly by dynamically associating with different network resources and networking algorithms; (ii) *cooperativeness*, i.e., SDN nodes can be implemented and aggregated at data centers for joint control and optimization to enhance the global network performance; (iii) *virtualizability*, i.e., multiple virtual wireless networks can be created on a single SoftAir, each of which operates under its own independent network protocols with network resources allocated based on demand; (iv) *openness*, i.e., data plane elements (i.e., BSs and switches), regardless of the underlying forwarding technologies and vendors, have unified data/control interfaces, e.g., CPRI and OpenFlow [5,6], thus significantly simplifying the data plane monitoring and management; and (v) *visibility*, i.e., centralized controllers have a global view of the network status collected from BSs and switches. The above five properties provide functionalities that are essential to enable 5G systems to possess the following promising features: evolvability and adaptiveness, infrastructure-as-a-service, maximal spectral efficiency, convergence of heterogeneous networks, and low carbon footprints.

The rest of the paper is organized as follows. Section 2 provides the state-of-the-art W-SDN and NFV. Section 3 introduces SoftAir, the proposed SDN architecture with NFV for 5G cellular systems. Section 4 concludes this paper.

2. State-of-the-art wireless software-defined network (W-SDN) & network function virtualization (NFV)

In the literature, the software-defined architectures are well-studied in wired networks. For example, in data center networks and campus local area networks (LANs) [4,7,8], these architectures mainly support centralized and adaptive manipulation of flow tables at switches and routers. Furthermore, considering wired network virtualization, cloud computing and computer virtualization have maintained strong foothold for the past few years. In particular, the virtualization of routers and switches has been adopted, such as virtual private networks (VPNs) over wide area networks (WANs) and metropolitan area network (MANs) as well as virtual LANs in enterprise networks. This is achieved by logically partitioning a physical network into virtual networks that share the physical routers/switches/crossconnects, physical links, and bandwidth on each link. The utilization of the physical resources needs to be carefully managed to maintain the QoS and security needs of the users of each virtual network. However, SDN's effectiveness and great potential for 5G data networking come with many new technical challenges, which need to be addressed by the new research advances.

2.1. Major problems with the current cellular architectures

Fig. 1 shows the current LTE network architecture and the corresponding data plane, which includes three components: cellular RAN, cellular CN, and the Internet. In particular, user equipment (UE) connects to eNodeB, i.e., base stations, and directs traffic through serving-gateway (S-GW) over a GPRS Tunneling Protocol (GTP) tunnel. S-GW serves as a local mobility anchor that enables seamless communication when the user moves from one BS to another. Towards this, S-GW must handle frequent changes in users' location, and store a large amount of user states since users retain their

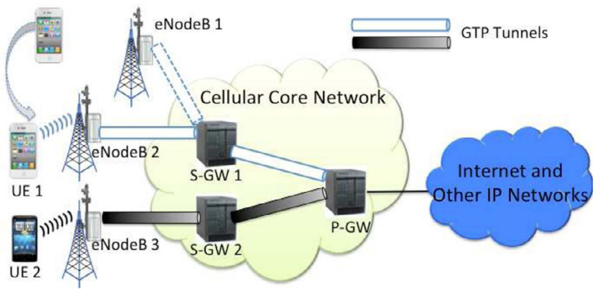


Fig. 1. LTE network architecture and data plane [9].

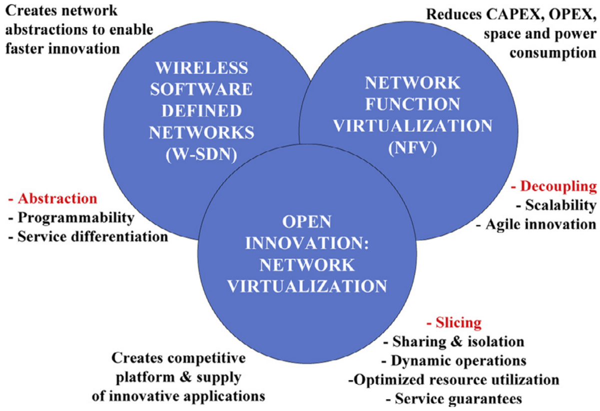


Fig. 2. W-SDN relationship with NFV and an example of open innovation, e.g., network virtualization.

IP addresses when they move. In addition, S-GW tunnels traffic to the packet-gateway (P-GW) which enforces QoS policies and monitors traffic to perform billing. P-GW also connects to the Internet and other cellular data networks, and acts as a firewall that blocks unwanted traffic. The bottom-line here is that besides data plane functionalities, eNodeBs, S-GWs, P-GWs also participate in several control plane protocols. Specifically, with mobility management entity, these devices perform hop-by-hop signaling to handle session setup, tear-down & reconfiguration, and mobility, e.g., registration, paging, and handoff. Moreover, S-GW and P-GW are also involved in routing such as OSPF. The policy control and charging function (PCRF) manages flow-based charging in the P-GW. It also provides the QoS authorization that decides how to treat each traffic flow, based on the user profile. The home subscriber server (HSS) contains subscription information for each user In case of cell congestions, a BS in coordination with P-GW reduces the max rate.

The major problems with the current cellular architectures lie as follows: (1) **Scalability challenges** and (2) **Vendor-specific device configuration**. First, centralizing data-plane functions such as monitoring, access control, and QoS functionality at P-GW introduces scalability challenges. It cause the equipment becomes very expensive, e.g., more than 6M for a Cisco P-GW. Moreover, centralizing data-plane functions at the cellular-Internet boundary forces all traffic through the P-GW. It becomes difficult to host popular content inside the cellular network. Second, network equipment has vendor-specific configuration interfaces, and communicate through complex control-plane protocols, with a large and growing number of tunable parameters (alone several thousands parameters for base stations). Hence, carriers (operators) have (at best) indirect control over the operation of their networks, with little ability to create innovative services. In short, existing commercial cellular systems rely on closed and inflexible hardware-based architectures both at the radio frontend and in the CN. These problems significantly delay the adoption and deployment of new standards, impose significant challenges in implementing and innovation of new techniques to maximize the network capacity and accordingly the coverage, and prevent provisioning of truly-differentiated services which are able to adapt to growing and uneven and highly variable traffic patterns. To tackle these problems, in the following, we summarize the design of flexible network architectures for 5G cellular systems, which are realized by the SDN paradigm with NFV.

2.2. Trends to W-SDN and NFV

As 5G cellular systems will be driven by software, the new architecture solutions heavily rely on two emerging technologies, i.e., SDN and NFV as shown in Fig. 2. In particular, the benefit of SDN lies in its ability to provide an abstraction of the physical network infrastructure. Through network-wide programmability, the capability to change the behavior of the network as a whole, SDN greatly simplifies the management of networks. The level of network programmability provided by SDN allows several network slices to be customized and optimized for different service deployments, while using the same physical and logical network infrastructure. Furthermore, by separating network function from underlying hardware devices, NFV allows a network function to be implemented in software either locally or on the remote clouds. This capability can enhance network scalability, which allows the optimal organization and easy management of network control and monitoring tools over the whole network. The most significant benefit brought about by NFV is the flexibility to execute and improve network management functions timely and independently of the underlying physical network forwarding infrastructure.

Recently, these concepts are realized in wireless cellular networks through different aspects [10–13], and the brief overviews with qualitative evaluation are investigated in the following.

(1) OpenRoads [10]: Fig. 3 shows an overview of OpenFlow wireless architecture by OpenRoads [10]. The main idea of this design is to separate the network service from the underlaying physical infrastructure and allow rapid innovation of network services. In particular, OpenRoads uses OpenFlow to separate control from the data path through an open API, FlowVisor [14] is applied to create network slices and provide isolation among them. Hence, it allows multiple experiments to run simultaneously in production wireless network. The SNMPVisor is to mediate device configuration access among experiments. These components virtualize the underlying infrastructure, in terms of decoupling mobility from physical network (OpenFlow from SDN/NFV), and allowing multiple service provides to concurrently control (FlowVisor from network virtualization) and configure (SNMPVisor from network

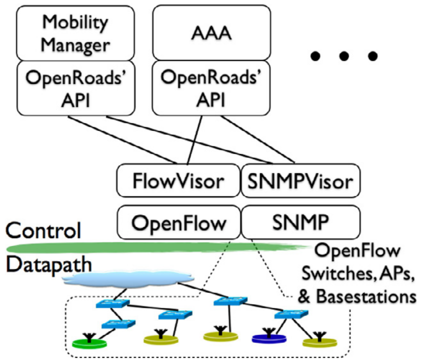


Fig. 3. OpenRoads: OpenFlow wireless architecture [10].

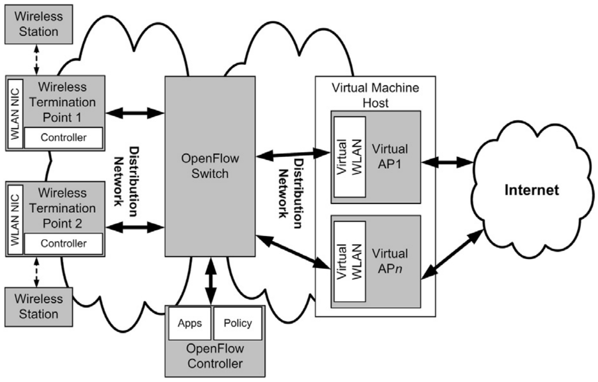


Fig. 5. Architecture of a CloudMAC based WLAN [12].

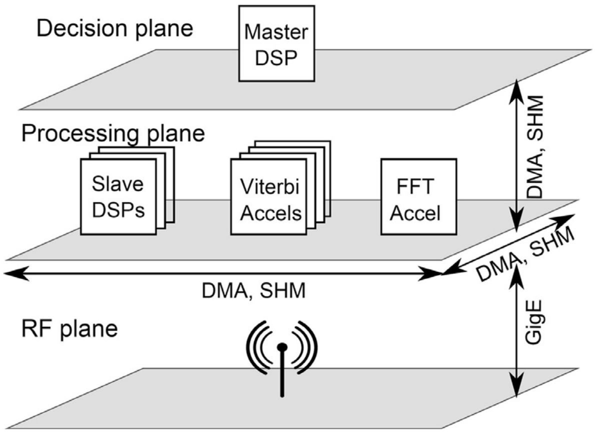


Fig. 4. OpenRadio BS internals [11].

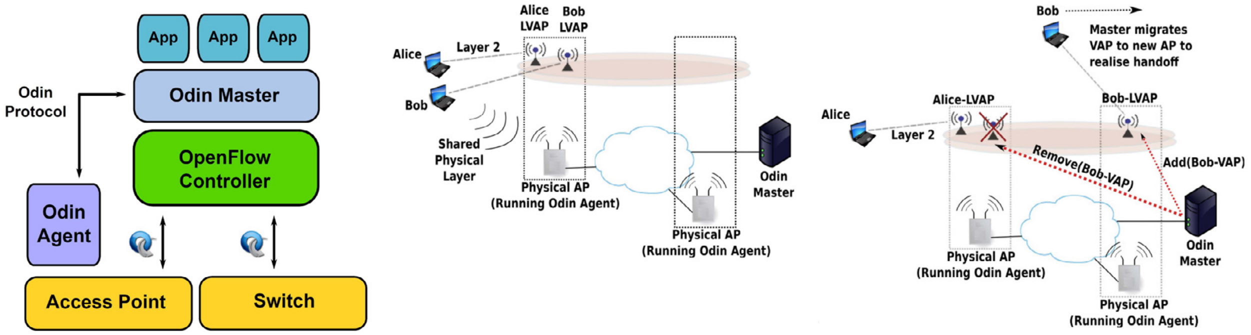
virtualization) the underlying infrastructure. The results are built and deployed as a backward compatible wireless network infrastructure for WiMAX service in collage campuses.

Limitations of OpenRoads: OpenRoads is the first work to move the wireless network forward on a path to greater openness. However, it mainly targets at WiMAX/WiFi networks with little support for cellular networks.

(2) OpenRadio [11]: A novel programmable wireless data plane is proposed by OpenRadio [11] that provides modular programming capability for the entire wireless stack. Fig. 4 shows the OpenRadio base station architecture, which focuses on utilizing multi-core DSP architecture. These architectures typically contain DSP cores optimized for signal processing computation and hardware accelerators providing speed-up of specific, commonly used algorithms. Such platforms provide the desired tradeoff between performance and flexibility. The software challenge is to harness the raw compute horsepower while retaining modular and declarative abstractions. On the other hand, the core objective of the runtime system, i.e., DSP cores and hardware accelerators, is fast and deterministic execution of actions. Towards this, while efficient execution is achieved through accelerated hardware, determinism in runtime is obtained through decision/processing separation down to the software level, mimicking the programming model itself. The same separation is applied to hardware resources as well, designating the decision plane core as the master core controlling the processing plane slave cores and accelerators. The decision plane runtime emulates the protocol state-machine by evaluating rules as transitions are made between states. The action associated with each state is computed on the processing plane by assigning constituent blocks to slave cores (or accelerators) according to a pre-computed execution and resource allocation schedule. Two major contributions of OpenRadio's design are as follows. First, it decouples wireless protocol definition from the hardware, even while ensuring that commodity multi-core platforms can be used for implementing the protocols. Second, the design of software abstraction layer that exposes a modular and declarative interface to program wireless protocols.

Limitations of OpenRadio: While OpenRadio introduces a novel design for a programmable wireless data plane, it does not define network controllers that can leverage the flexibility offered by a programmable data plane. In other words, OpenRadio focuses more on the NFV, rather than the concept of W-SDN for cellular systems.

(3) CloudMAC [12]: CloudMAC [12] is a network architecture aimed at having a programmable MAC layer without resorting to software radios. It is a distributed architecture in which 802.11 WLAN MAC processing is partially performed in data centers on virtual machines connected by an OpenFlow controlled network. As shown in Fig. 5, a CloudMAC based WLAN deployment consists of virtual APs (VAPs), wireless termination points (WTPs), an OpenFlow switch, an OpenFlow controller and tunnels to connect the entities. In particular, VAPs are operating system instances running on a virtualization host, such as Xen or VSphere Center. Each VAP has one or several virtual WLAN cards. Moreover, WTPs are slim APs with WLAN cards that allow to send and receive raw MAC frames. Specifically, each WTP is *dumb* in that they do not generate their own MAC frames. WTPs are used only to send and receive MAC frames from an OpenFlow switch that they are connected to, in addition to the end user. VAPs and WTPs are connected with each other via MAC layer tunnels and an OpenFlow switch. The switch contains a switch table, which specifies what frames to forward to which WTP or VAP. The controller runs applications that configure the switch table using the OpenFlow protocol. The forwarding table represents the binding between VAPs (more specifically, the virtual WLAN cards) and WTPs. Towards thus,



(a) Architecture of Odin. (b) Light virtual AP (LVAP) abstraction.

Fig. 6. Odin [13].

CloudMAC enables several new applications, such as dynamic spectrum use, on-demand AP, and downlink scheduling. Moreover, the testbed evaluation shows that CloudMAC achieves seamlessly interworks with standard IEEE 802.11 stations, increasing WLAN flexibility while at the same time largely reducing the CAPEX and OPEX accordingly.

Limitations of CloudMAC: A shortcoming of CloudMAC is the potential increased packet delay. The reason is that all MAC frames generated by the user will have to travel to the VAPs which could potentially lie somewhere deep in the network. It causes poor delay and indeed their implementation shows a three fold increase in round trip time.

(4) Odin [13]: Odin [13] in Fig. 6 is an SDN framework that proposes to simplify the implementation of high level enterprise WLAN services, such as authentication, authorization and accounting (AAA), by introducing light virtual access points (LVAPs). This approach is similar to the virtual access points used in CloudMAC. As shown in Fig. 6a, Odin's architecture consists of a single master, multiple agents and a set of applications. The Odin master itself is an OpenFlow application on top of an OpenFlow controller. It speaks OpenFlow to the switches and the APs and uses a custom protocol to talk to each Odin agent running on APs. A TCP connection is used between the agent and the master, which serves as Odin's control channel to invoke commands on the agents and collect statistics from them.

Usually, the AP, which a user is connected to, may change in accordance with local decisions made by the user. Thus the last hop connecting the user to the WLAN infrastructure is not stable. Using the abstraction of LVAP, Odin gives programmers a virtual unchanging link connecting users to APs. To achieve this each user is assigned a unique BSSID, giving the illusion that it has its own AP. This virtual user-specific AP is called a LVAP. Each physical AP will host multiple LVAPs, one corresponding to each client. For example, in Fig. 6b, at first, clients connect to LVAPs, with several LVAPs being hosted on the same physical AP. Then, the Odin master detects a client movement, i.e., Bob, and performs an LVAP migration to realize a smooth handoff. The advantages of moving to this centralized Odin architecture, such as easy load balancing and mitigating the hidden terminal problem, are preserved. It is also easier to implement mobility managers since the BSSID at the user does not change during a handover.

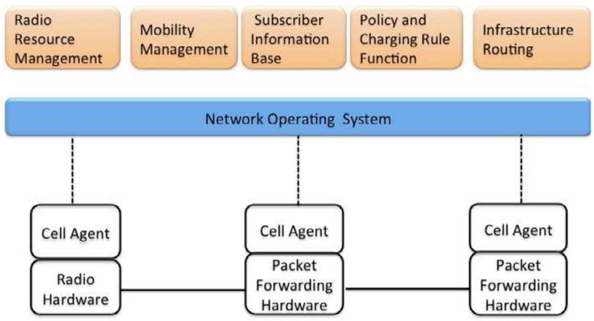


Fig. 7. Cellular SDN (CellSDN) [9].

Limitations of Odin: Similar to CloudMAC, Odin might also has delay issue. As one moves more and more intelligence to the center of the network, it might not be acceptable as centralized controller architecture of SDN might prove too slow in reacting to PHY or MAC layer incidents. The optimal operating point depends on the specific network requirements. Note that OpenRoads [10], OpenRadio [11], CloudMAC [12], and Odin [13] are all partial solutions that exploit one or many concepts of W-SDN and NFV (even network virtualization) as in Fig. 2, and do not provide a completed/integrated architecture/plateform for the 5G cellular systems.

2.3. Integrated W-SDN & NFV solutions

In this section, the fully integrated W-SDN and NFV solutions (i.e., [9,15–20]) are introduced for practical implementations of 5G cellular systems.

(1) CellSDN [9]: Fig. 7 summarizes the network architecture of cellular SDN (CellSDN), which aims to achieve a centralized control plane for cellular CNs. SDN architecture offers fine-grain, real-time control without sacrificing scalability. It enables SDN in cellular networks with 4 more extensions: (1) controller applications express policy in terms of subscriber attributes; (2) switches run local control agent for simple actions; (3) switches support more flexible data

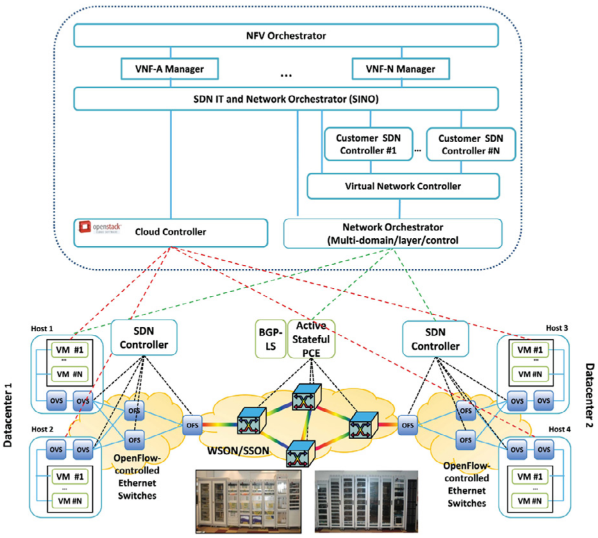


Fig. 8. SDN/NFV cloud computing platform and CN by ADRENALINE [15].

plane functionality; (4) BSs support remote control of virtualized wireless resources. SDN controller consists of network operating system (NOS) running many application modules, e.g., radio resource management, mobility management, and routing. Also, SIB stores and maintains subscriber information, including static subscriber attributes, such as dynamic data like users' current IP address, location and total traffic consumption. NOS can translate policies into switch rules, such as matching decisions with respect to packet headers or to network measurements to the subscribers, allowing application modules to focus on subscribers and their attributes. Controller can also dynamically divide the network into *slices* that handle all traffic matching some predicate on the subscriber attributes. It allows the provider to isolate roaming traffic or phone traffic using legacy protocols.

In addition, controller may not be able to respond as fast to local events as the underlying switches themselves. Switches have some simple control software (called local control agent) performing simple local actions, under the command of the controller, e.g., automatically changing the weight or priority of a queue when traffic exceeds a threshold,or pushing a tag on a packet to direct the traffic through an intermediate middlebox. Hence, the research challenge for CellSDN is to design of the local agents and techniques for partitioning functionality are subject to research.

Limitations of CellSDN: Since CellSDN is the first work to incorporate SDN into cellular networks, it only for cellular CN and has no consideration about RAN. It provides an abstract concept without practical realization methods. In particular, a premature decoupling architecture is introduced, including (i) controller and switches with local control agents, (ii) switches with flexible data plane functionalities, such as deep packet inspection and header compression. CellSDN also mentions the basic virtualization functionalities that BSs, at the behest of controller, should support virtualization of wireless resources for flexible cell management.

(2) ADRENALINE [15]: ADRENALINE [15] provides a cloud computing platform as shown in Fig. 8 that serves as an industrial solution of SD-CN for 5G cellular systems with optical OFDM (O-OFDM). It includes a multi-technology control plane for multilayer (packet over optical) networks, which manages the networking resources and covers the long-haul core transport and aggregation segments. The design of the ADRENALINE control plane follows broad SDN principles, such as stacking components in a hierarchical setting with different levels of abstraction. Network connectivity services are provisioned by an overarching control orchestration. In particular, at a given domain and layer, the control plane can be based on the GMPLS technology and protocols, a distributed system in which a dedicated controller is responsible for each node autonomously, or follow SDN/OpenFlow principles, with a centralized controller that manages all the aspects of a network, dynamically configuring networks according to users an application needs. In addition, the design of flexible, O-OFDM transmission schemes is studied and assessed through an experimental platform. The experimental validation of the investigated modulation schemes based on multicarrier technology is enabled and comprises two different parts. One is a photonic mesh network of 4 nodes with links of up to 150 km that is used to experimentally assess the different optical OFDM systems. The other one encompasses the different setups, including optical and optoelectronic systems and subsystems, for multicarrier modulation (either OFDM or DMT) transmission and reception based on offline processing.

End-to-end network orchestration (to provide an overarching control regardless of the number of domains) is also enabled with extensive usage of the application-based network operation architecture and framework, using the services of the ADRENALINE control plane. End-to-end network virtualization services are performed by a virtual network controller, which is able to provide abstracted multi-layer network views to customers, ensuring security, isolation and independent SDN control (i.e., customer SDN controllers). In particular, ADRENALINE includes an SDN integrated IT and network orchestrator (SINO), which is a centralized system able to coordinate, from a high-level view, cloud and network service management aspects in modern multi-tenant environments, which provide the platform to run user applications and virtualized network functions (VNF Manager). A NFV orchestrator is also provided in to deploy end-to-end VNF through VNF Forwarding Graphs. The cloud computing service manager is implemented in terms of a modified OpenStack software [21], one of the top open-source distributed cloud computing systems.

Limitations of ADRENALINE: ADRENALINE only provides some ideas for CN without the incorporation of RAN. Also no scalability design for SD-CN is included and limited SD solutions in SD-CN, i.e., optical OFDM systems, is considered. Note that CellSDN and ADRENALINE neither consider the scalability issue of SD-CN nor the incorporation of CN with RAN. On the other hand, SoftRAN [16] attempts to restructure the control plane of RANs in a software-defined manner as follows.

(3) SoftRAN [16]: While current LTE's distributed control plane is suboptimal in achieving wireless connectivity through limited spectrum management, SoftRAN [16] introduces a software-defined centralized control plane for RANs that abstracts all BSs in a local geographical area as a virtual big-base station comprised of a central controller and radio

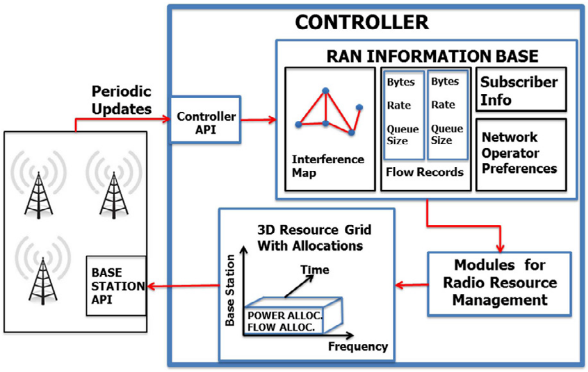


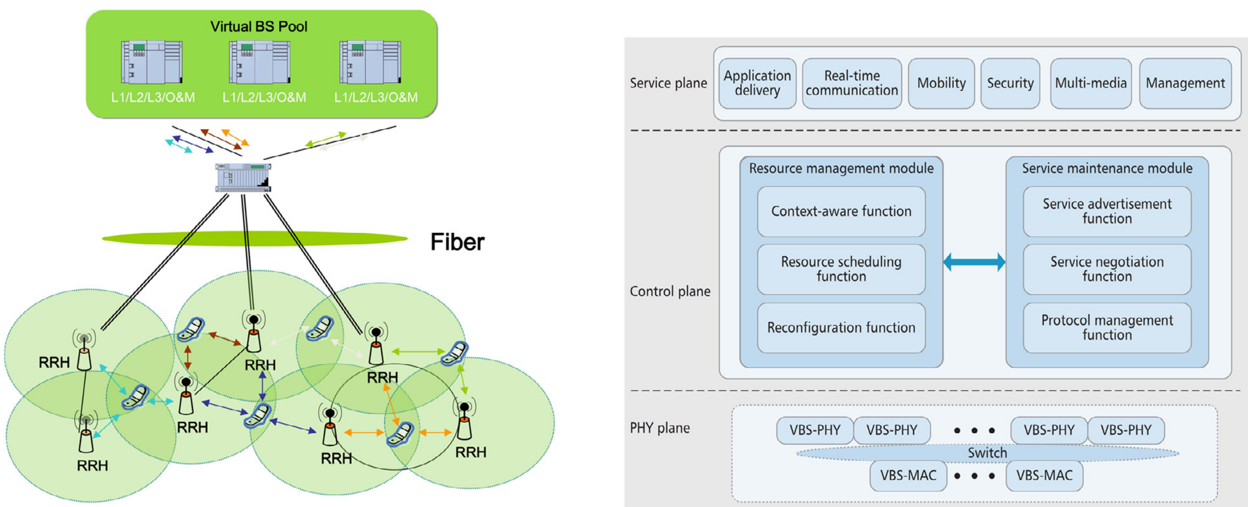
Fig. 9. SoftRAN architecture [16].

elements. In particular, the SoftRAN architecture, as shown in Fig. 9, aims to (i) design a controller which can provide a framework for different control algorithms to operate on, and (ii) to ensure that the delay between the controller and the radio element does not negatively impact performance. Towards this, a centralized controller is deployed, which receives periodic updates of local network state from all the radio elements in a local geographical area. Given these updates, the controller updates and maintains the global network state in the form of a database, called RAN information base (RIB). The RIB conceptually consists of the following elements: Interference map, flow Records, and network operator preferences. Through this abstraction and architecture, an environment is created that enables efficient and dynamic management of increasingly scarce and strained radio resources.

Limitations of SoftRAN: SoftRAN builds up the centralized control plane for SD-RAN. However, it addresses limited NFV over the antennas in BSs for big-base station abstraction. Also, more detailed consideration about the interaction with SD-CN is missing in this SD-RAN design.

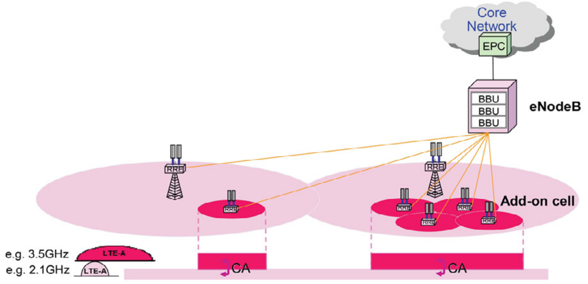
(4) Cloud-RAN (C-RAN) [17,22]: An emerging distributed RANs, called Cloud-RAN [17,22] in Fig. 10, proposes SD-RAN architecture that connects SD-RANs to virtual BS pool through fibers and provides centralized control solution upon the BS pool. This solution is widely employed by many Telcom companies, such as IBM, Intel, Huawei, ZTE. Fig. 10a summarizes the network architecture of C-RAN, which is composed of the baseband unit (BBU), optical transmission network (OTN), and remote radio head (RRH). In particular, distributed radio units referred to as RRHs plus antennas which are located at the remote site; high bandwidth low-latency optical transport network which connect the RRHs and BBU pool; BBU composed of high-performance programmable processors and real-time virtualization technology. Their functionalities are explained as follows. All BS computational resources are aggregated into a central pool and radio signals from different antennas are collected by RRHs and transmitted to the CLOUD platform through an OTN. Specifically, BBU acts as a digital unit implementing the BS and RRHs perform radio functions, including frequency conversion, amplification, and A/D and D/A conversion. Moreover, RRHs send/receive digitalized signals to/from the BBU pool via optical fiber. Antennas are equipped with RRHs to tx/rx RF signals. By placing numerous BBUs in a central physical pool while distributing RRHs according to RF strategies, operators can dynamically employ a real-time virtualization technology that maps radio signals from/to one RRH to any BBU processing entity in the pool.

Consider the service cloud structure of C-RAN as shown in Fig. 10b. For the physical plane, C-RAN can use general-purpose processors (GPPs) with multicore and multithread techniques to implement virtualized and centralized baseband and protocol processing, such as PHY and MAC layers. This PHY deals with the virtualization for resource provisioning, baseband pool interconnection, and signal processing. Second, for the control plane, C-RAN consists of two module: resource management module and service

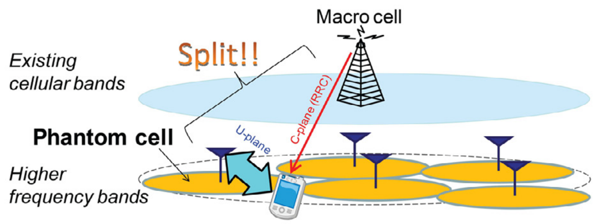


(a) SD-RAN. [10] (b) Service cloud structure. [31]

Fig. 10. Cloud-RAN (C-RAN). [17,22].



(a) DOCOMO'S architecture.



(b) Phantom cell with C/U-plane splitting.

Fig. 11. Advanced C-RAN architecture with phantom cell from NFV by DOCOMO [18].

maintenance module. In particular, the former takes in charge of available radio and computational resources from network and users perspectives to realize high QoS, seamless mobility, and power utilization efficiency. On the other hand, the maintenance module takes in charge of available services from the perspective of the network, as well as negotiation and realization of services between network providers and users. Third, for the service plane, a platform where fixed and mobile services are provided and managed by TELCOs. Subscribers obtain services from the cloud as if it is a black box, while each service can be supported by multiple radio access technologies (RATs) simultaneously. This plane comprises a scalable library of NW-based services to deliver voice, data, and multimedia applications in a consistent, robust, and efficient manner. In short, C-RAN has reduced cost for saving on CAPEX/OPEX, better energy efficiency and green infrastructure, improved spectrum utilization, adaptability to non-uniform traffic, smart Internet traffic offload, and business model transformation. Hence, the research challenge for C-RAN is a radio over low cost optical network, advanced cooperative transmitter and receiver, baseband pool interconnection, BS virtualization technology, and service on edge.

Limitations of C-RAN: Although providing a practical solution for wireless systems than CellSDN, C-RAN has limited scalability from its coarse-grained BS decoupling. In particular, all PHY layer functions located at remote baseband processing servers (BBS) or virtual baseband pool, where I-Q plane data must be exchanged between BBS and RRH. In other words, C-RAN keeps MODEM functionality in far cloud for both fully and partial centralized architectures. Towards this, data rate required to transfer I-Q plane data linearly proportional to radio BW, number of antennas, and duplex mode. For example, transporting a 20 MHz LTE waveform between BBS and a half-duplex RRH with 8 antennae, a transmission rate of 7.36 Gbps is required to send the digital I-Q plane data. This brings huge burden on the fronthaul network as the current trend of wireless technologies moving towards massive MIMO, full-duplex transceivers, mm-waves, TeraHertz band, thus significantly degrading the scalability and evolvability of the current distributed RANs. C-RAN also has limited exploration of virtualization functionalities and only examines SD-RAN without the incorporation of CN.

(5) DOCOMO [18]: As shown in Fig. 11, DOCOMO exploits C-RAN concept in 5G RAN design to split Control (C) and User Data (U) Plane, and to move control functionality in far-end cloud for centralized network architecture with remote radio equipment. It further adopts carrier aggregation (CA) and small-cell technologies, bringing advanced C-RAN architecture as shown in Fig. 11a. In particular, small cells with low power nodes cope with mobile traffic explosion. And the CA functionality is adopted between macro and small-cell carriers. Specifically, it maintains basic connectivity and mobility under macro-Cell coverage and small-Cells (called add-on cells) achieve higher throughput and larger capacity. In addition, the advanced C-RAN architecture handles all processing for CA and handover within a centralized baseband unit at eNodeB, which drastically reducing signaling to CN.

One of DOCOMO's major contribution is to investigate phantom cell scenario from NFV in Fig. 11b. In particular, phantom cell, based on multi-layer network architecture, splits control (C)-plane and the user data (U)-plane between macro-cell and small-cells using different frequency bands. DOCOMO's 5G concept uses the Phantom cell architecture and integrates such multi-layer NWs using lower and higher frequency bands. Specifically, small-cell handles traffic for high-throughput data sessions with user (U-plane); macro cellular layer controls C-plane signaling, e.g., radio resource control (RRC). Macro- and small-cells form master-slave relationship, through which macro-Cell sends control info to user connected to small cells, i.e., make small-cells practically invisible to user from the name of phantom cell.

Limitations of DOCOMO's Advanced C-RAN: By proposing an advanced C-RAN architecture, DOCOMO focuses on the decoupling and NFV for higher frequency bands in realizing phantom cells. However, it still suffer from tremendous I-Q transmissions, like C-RAN, which limits the scalability. In addition, it does not consider much for network virtualization. In particular, it suggests non-orthogonal multiple access (NOMA) to further enhance cellular systems in lower frequency bands. Coherent (resource) slicing scheme is still needed to realize the core virtualization idea of SDN into wireless cellular systems.

(6) SK Telecom [19]: SK Telecom introduces a software-based 5G enabling platform as shown in Fig. 12, which is an industrial solution for 5G cellular systems that provide software-oriented framework and Telco asset-based interface, while jointly considering SD-RAN and SD-CN. It decouples control (software functionalities) and data plane (hardware devices), and a core function in software framework is to provide network-as-a-service platform, which allows configuration/change of telecommunication and service

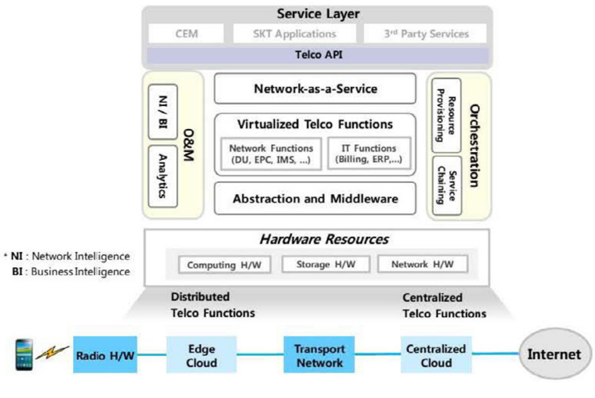


Fig. 12. Software-based 5G platform by SK Telecom [19].

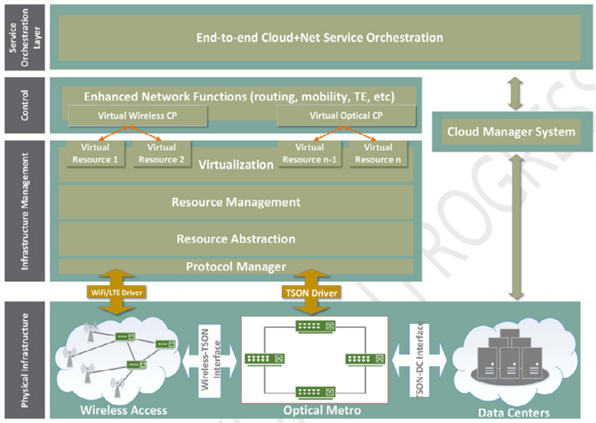


Fig. 13. CONTENT layered architecture [20].

functions. SK Telecom also provides Telco API for service utilization, enabling implementation of analytics-based services, i.e., multi-Service carrier Ethernet 2.0 and MPLS edge solution. The proposed architecture jointly realizes SD-CN and SD-RAN, where SD-CN includes programmable dumb switches with flow tables, SD-RAN has radio hardware units and antennas, and a single coherent control plane is applied. Towards this, SK Telecom realizes many promising 5G technologies through this software-based enabling platform.

Three important technologies are supported from SK Telecom's evolution of software-based network through NFV and SDN: NFV-based virtualization CN operation, virtualized RAN, and SDN and integrated orchestration. First, via NFV, SK Telecom builds the cloud by virtualizing a standard hardware and operate a range of network/service functions on the software-based network. Second, SK Telecom enables a technology to centralize and virtualize digital unit of a BS into a standard hardware-based cloud and process RAN signals in real time. Third, it provides an effective control and life-cycle management of the software-based networks services from a centralized & unified network service orchestrator.

Limitations of SK Telecom's Software-Based 5G Enabling Platform [19]: While providing a practical solution to consider both SD-CN and SD-RAN, SK Telecom's solution has limited scalability due to coarse-grained BS decoupling as C-RAN and control traffic unbalancing. In particular, SK Telecom's solution transfer from hardware-oriented design into software one for realizing control/data decoupling. However, it does not consider in details for control and data plane decoupling and still suffers from tremendous I-Q transmissions. Also, no coherent scheme to realize network virtualization exists in SK Telecom's solution, where wireless resource slicing scheme is missing and thus cannot provide efficient and flexible cell management.

(7) CONTENT [20]: CONTENT (Convergence of Wireless Optical Network and IT Resources in Support of Cloud Services) [20] is a 3-years European co-funded (FP7) project, which started in November 2012 and will end in October 2015. It aims at offering a network architecture and overall infrastructure solution to facilitate the deployment of conventional cloud computing as well as mobile cloud computing, which can introduce new business models and facilitate new opportunities for a variety of business sectors. Fig. 13 shows the layered architecture proposed by CONTENT, which includes several layers: (i) heterogeneous physical infrastructure layer, (ii) infrastructure management layer, (iii) control layer, and (iv) service orchestration layer. The details are briefly introduced as follows. First, heterogeneous physical layer includes a hybrid wireless access network (LTE/Wi-Fi) domain, and an optical metro network domain (TSON) interconnecting geographically distributed data centres, supporting frame-based sub-wavelength switching granularity. Second, infrastructure management layer manages the creation of virtual network infrastructures over the underlying physical resources. An important feature of the functionalities supported, is orchestrated abstraction of resources across domains, involving information exchange and coordination across domains. Third, control layer establishes seamless connectivity across heterogeneous technology domains (wireless access and optical metro) through a coordinated, end-to-end approach to support optimized performance, QoS guarantees as well as resource efficiency and sustainability. Last, the orchestration layer is responsible for efficient coordination of the cloud and network resources, in order to enable the end-to-end composition and delivery of integrated cloud, mobile cloud and network services in mobile environments supporting the required QoE. In short, CONTENT proposes a layered architecture with the aim to facilitate the main principles of its novel proposition i.e. cross-technology virtualization in support of optimized, seamless and coordinated cloud and mobile cloud service provisioning across heterogeneous network domains.

Limitations of CONTENT: While CONTENT has the most completed solution that covers the concept of W-SDN, NFV, and network virtualization so far, more detailed functionalities supporting this solution are missing. In other words, the entire framework covers all essential ideas for 5G cellular systems, but more deep consideration and concrete mechanisms are still mising.

In sum, DOCOMO [18], SK Telecom [19], and CONTENT [20] all provide their own industrial SD-RAN solution for 5G systems, respectively. Similar to Cloud-RAN, these solutions adopt coarse-grained decoupling and thus, bring bottlenecks to fronthaul fiber networks. Moreover, different from our proposed W-SDN architecture, i.e., SoftAir [3] in Section 3, all of the above architectures lack a coherent framework that integrates cellular CN and RAN in a software-defined

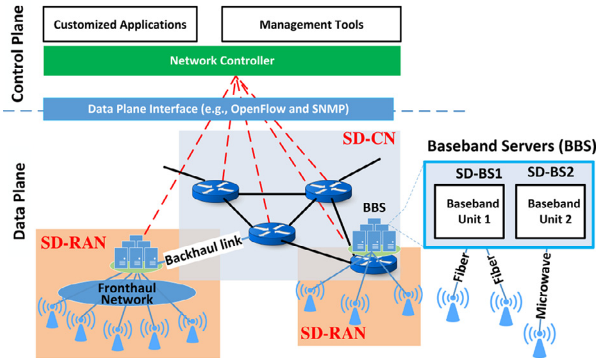


Fig. 14. Overall architecture of SoftAir [3].

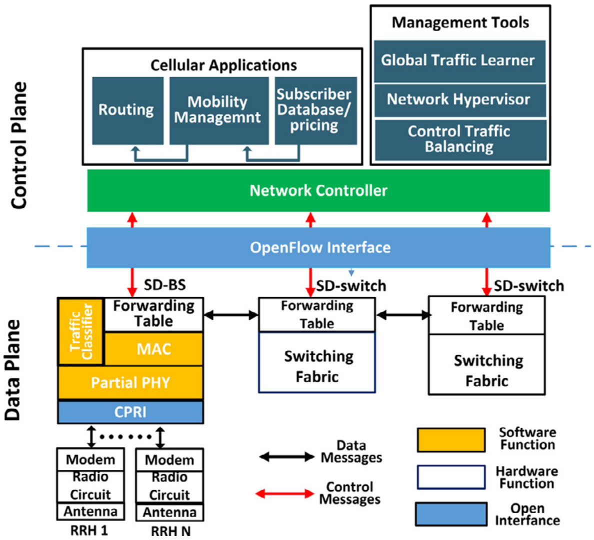


Fig. 15. NFV of SoftAir [3].

manner. What is more important, these solutions lack underlying concrete algorithms that leverage the promising properties of software-defined wireless network architecture. In addition, network virtualization, serving as a key building block to enable the sharing of (mobile) carrier networks for enhanced spatio-temporal spectrum utilization [23], is omitted by all existing software-defined architectures.

3. SoftAir: A software-defined networking architecture for 5G cellular systems

SoftAir consists of a data plane and a control plane as shown in Fig. 14. The data plane is an open, programmable, and virtualizable *network forwarding infrastructure*, which consists of SD-RAN and SD-CN. The SD-RAN consists of a set of SD-BSs, while the SD-CN is composed of a collection of SD-switches. The *control plane* mainly consists of two components: (1) network management tools, and (2) customized applications of service providers or virtual network operators. In the following, we present the scalable SoftAir architecture in detail, explain the network virtualization, and introduce three essential management tools, namely mobility-aware control traffic balancing, resource-efficient network virtualization, and distributed/collaborative traffic classifier.

3.1. Network function virtualization (NFV)/scalable network function cloudification

NFV, also known as network function cloudification, decouples network functions from the underlying hardware and centralizing/cloudifing them at network servers, making the network architecture highly flexible from quick and adaptive reconfiguration. Indeed, the main advantage of the wired OpenFlow-based SDN relies on the network-layer function virtualization, which decouples the routing function from the hardware switches and centralizes it at a network controller through an open network interface of OpenFlow. Despite its great advantages, such cloud-based network architecture also imposes challenges on the network scalability.

(1) Scalable SD-CN: To provide the cellular CN with high flexibility, our proposed SoftAir adopts SD-switches to form the SD-CNs as shown in Figs. 14 and 15. By SD-CNs, the customized SDN applications, e.g., mobility management, QoS routing, and billing policies; as well as global management tools, e.g., traffic classification and network virtualization, can be designed, deployed and updated on the network controller to fit the specific and ever-changing needs. The practicality of SD-CN can be envisioned because of the successful field deployment of SDN, e.g., B4 from Google [2], SWAN from Microsoft [24], and ADMCF from Huawei [25]. The recent deployments of software-defined WANs, e.g., B4 from Google [2], has successfully demonstrated the promising performance of SDN by boosting the average link utilization from 30–40% to over 70%. The scalability of SD-CN is greatly improved by leveraging high-performance controllers and optimized network management. For example, with the current SDN technology, one single controller can achieve 12 million requests per second processing speed for the control messages between the switches and the controller [1]. Such high processing capacity has been further enhanced by the adoption of controller clusters and advances in multi-threading technologies [1]. More importantly, our recent research [26] has shown that even in large-scale SDN with practical in-band control channels, the control message forwarding delay between the controllers and the switches can be minimized by our proposed control traffic balancing scheme, which employs emerging parallel optimization theories to achieve fast and reliable control message forwarding. To further increase the scalability of SD-CN, we can adopt a *mobility-aware and proactive control traffic balancing scheme*, which minimizes the control message forwarding delay by taking into account the unique mobile feature of SD-RANs (formally defined in the following subsection), where the control traffic oriented from SD-RAN are dynamic but follows certain spatial and temporal patterns. Accordingly, this balancing solution can employs a hybrid approach.

(2) Scalable SD-RAN: To further enhance the network flexibility, we simultaneously realize the physical-, MAC-, and network-layer function virtualization for SD-RAN. As in Fig. 14, the proposed SD-RAN follows a distributed RAN architecture. Here, each SD-BS is split into hardware-only radio head(s) and software-implemented baseband units (these

two components are not necessarily co-located). In particular RRHs are connected to the baseband units on BBS through fronthaul network (fiber or microwave, which depends on available infrastructure) using standardized interfaces, such as CPRI or OBSAI interface. The distributed RAN architecture has recently received significant attention from both industry and academia. C-RAN [17] is an example. The applicability and practicality of distributed RAN is enabled by the fast evolution of software-defined radio architectures. Significant efforts have been devoted to the standardization of open and technology-independent communication interfaces to connect RRHs and BBS. The standardized interface CPRI, jointly developed by Ericsson, Huawei, NEC, Alcatel Lucent and Nokia Siemens, can enable high-speed (up to 10 Gbit/s, low bit error rate (10^{-12}), and long-distance (up to 40 miles) data exchange between RRHs and BBS, while providing the high-resolution synchronization. Existing distributed RAN architectures such as C-RAN mainly focus on the high-performance computing of baseband processing functions (mostly for physical layer operations) at remote servers or data centers. It faces two fundamental limitations. First, they cannot achieve scalable PHY/MAC-layer cloudification. Second, they do not support network-layer cloudification as the SD-CN. To this end, our proposed SD-RAN offers significantly enhanced scalability, evolvability, and cooperativeness through fine-grained BS decomposition that overcomes fronthaul traffic burden.

Fine-Grained BS Decomposition. The existing distributed RAN adopts a coarse-grained function splitting between RRHs and BBS, where the entire baseband processing along with MAC layer operations are located at the BBS, while RRHs only implement the radio frontend. Through this approach, the digital I-Q samples must be transported between BBS and RRHs, which inevitably demands extremely high data rates on the transport network (i.e., fronthaul network). More specifically, the data rate requirement R_{IQ} for transporting I-Q samples of a waveform depends on the bandwidth W, the oversampling ratio $r > 1$, the quantification resolution Q_r of the analog signal, the duplex mode D_m, and the number of antennas A at the radio head ($R_{IQ}(B, A, D_m) = 2 \times W \times r \times Q_r \times D_m \times A$[4]. To address this challenge, SoftAir adopts a new fine-grained BS decomposition architecture by leaving partial baseband processing at the RRH (e.g., modem), while implementing the remaining baseband functions, e.g., MIMO coding, source coding, and MAC, at the BBS, as shown in Fig. 15. Such function splitting is convenient because CPRI, which is not only defined for I-Q sample transport, can still be adopted without designing new interfaces and can lead to considerably reduced data rate requirements between BBS and RRHs. In addition, our decomposition still preserves sufficient flexibility offered by the distributed RAN architecture. In summary, with reduced data rate requirements, our SD-RAN offers excellent scalability, cooperative gain, and evolvability to next-generation wireless networks by allowing the aggregation of a large number of technology-evolving RRHs at BBS through the diverse, cost-efficient, CPRI-supported fronthaul network topologies, and over different fronthaul mediums.

[4] The factor of 2 captures both I-plane and Q-plane digital samples.

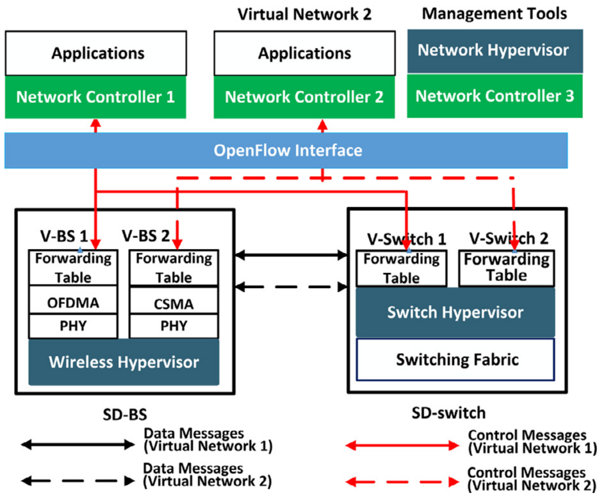


Fig. 16. Network virtualization of SoftAir [3].

Seamless Incorporation of OpenFlow. Different from the existing distributed RAN [17], our proposed SD-RAN incorporates OpenFlow into the SD-BS. In our SD-RAN design, we implement an OpenFlow interface for each SD-BS by utilizing Open vSwitch (OVS) [27], which is an OpenFlow-capable software switch that can easily be realized in BBS. With OVS, each SD-BS will be able to interpret, exchange, and respond to the OpenFlow protocol messages. Equipping SD-BSs with OpenFlow capabilities provides a unified interface to control and manage base stations with different wireless standards, thus leading to a multi-technology converged RAN that allows smooth transitions among different radio access technologies. For example, with OpenFlow-capable SD-BSs, seamless vertical mobility can be easily achieved, which allows mobile users to transparently roam among BSs with different wireless standards, without experiencing disruptions in network services. This can be done by simply re-routing the traffic flows through the CN over the best paths to different BSs via the technology-independent OpenFlow interface enabled on both core network switches and BSs. Moreover, adopting the common OpenFlow interface on both SD-switches and SD-BSs promises the transparent interconnections between SD-CNs and SD-RANs and allows the unified management of the entire SoftAir.

3.2. Network virtualization

Network virtualization enables multiple isolated virtual networks, e.g., M2M, smart grid, over-the-top service provider, cellular provider, to share the same physical network infrastructure, as shown in Fig. 16. It focuses on slicing network resources for multiple virtual networks so that they can simultaneously share the same physical network architecture. More specifically, each virtual network can adopt its customized PHY/MAC/NET layer protocols, without interrupting the operations and performance of other virtual networks. These virtual networks can also be deployed on demand and dynamically allocated. The proposed SofTAir with network virtualization capability enables a wide range of emerging applications, e.g. (1) allowing MVNOs to adopt different wireless standards (i.e.,small cells, HetNets, LTE, WiMAX, or WiFi), (2) enabling active RAN sharing, which could allow the operators to save up to $60 Billion in a period of 5 years, through the significant reduction in equipment investments in low traffic areas [28], (3) promising virtualization-enabled QoS routing by simultaneously satisfying the strict end-to-end performance requirements (e.g., delay, jitter, and throughput) of different network services that generate fundamentally different traffic flows [29], and (4) accelerating technology innovation by allocating isolated wireless resources to deploy and test innovative technologies on the operational networks in large-scale real-life scenarios. To realize these isolated virtual networks, SoftAir implements three functions: network hypervisor for high-level virtualization as well as wireless hypervisor and switch hypervisor for low-level virtualization.

(*1) Network Hypervisor*: It focuses on high-level resource management, which determines how to distribute non-conflicting network resource blocks among virtual network operators based on their demands. An utilization-optimal network hypervisor is proposed to maximize the global resource utilization, while guaranteeing the data rate requirements demanded by each virtual operators. Given a set of virtual operators, the data rate required by a virtual network can be formulated, which means that within the coverage area of each SD-BS, the virtual network needs to offer a certain average data rate for its users with certain spatial distribution and density. Then, at each SD-BS, we can assign the wireless resource blocks to the virtual network. With the allocated resource blocks, power, and RAT, the average data rate that this SD-BS can offer to the virtual network is thus obtained, where users follow certain spatial distribution. This average data rate relates to the effective SINR at an arbitrary user of the SD-BS and the distance vectors from each user to each RRH within the considered SD-BS.

(*2) Wireless Hypervisor:* It is a low-level resource scheduler that enforces or executes the resource management policies determined by the network hypervisor by employing a variety of wireless resource dimensioning schemes, e.g., OFDMA or wireless scheduling, so that 100% isolation among virtual networks is guaranteed. As a consequence, each virtual network can implement and adopt its own and customized NET/MAC/PHY layer protocols. Besides providing 100% isolation between the virtual networks, the wireless hypervisor has to ensure efficient utilization of the limited spectrum resources, which exhibit inherent channel and user diversity.

(*3) Switch Hypervisor:* Enabled by OpenFlow protocol, switch hypervisor (or switch fabric) focuses on bandwidth partitioning in a single SD-switch. In particular, bandwidth provisioning on switches aims to offer predefined bandwidth for any specific traffic flow or virtual network. An well-known FlowVisor is an layer responsible for the isolation among slices of the virtualized infrastructure and employs leaky-bucket scheme for bandwidth provisioning, which however, is not a work-conserving policy and thus can not guarantee full bandwidth utilization. To solve this issue, the recent work [30] exploits GPS models upon combined input cross-point queued (CICQ) switches, which ensures full bandwidth utilization while providing accurate bandwidth provision.

Remark. By utilizing the proposed wireless hypervisor at SD-BSs along with the Flowvisor at SD-switches, SoftAir will have the capability to enable the end-to-end network virtualization traversing both SD-RAN and SD-CN, thus realizing a truly multi-service converged network infrastructure.

3.3. SoftAir management tools

Cloud orchestration aims to automate the configuration, coordination and management of software and software interactions in the cloud environment. To support cloud orchestration in SoftAir as shown in Fig. 15, three essential and general management tools need be developed: (1) mobility-aware balancing, (2) resource-efficient virtualization, and (3) distributed and collaborative traffic classifier.

(*1) Mobility-Aware Control Traffic Balancing:* To promise on-line and adaptive traffic engineering in SoftAir, the control messages should be forwarded from SD-switches or SD-BSs to the controller(s) in a fast and robust manner. Moreover, due to mobile users, the unique mobile traffic features in SD-RANs should be examined for control traffic balancing. Inspired by our previous work [26], SoftAir provides a novel mobility-aware control traffic balancing that decides the optimal controller locations and optimal forwarding paths of control flows with respect to the mobile traffic. First, through the unique features of SoftAir, users' mobility pattern and the corresponding traffic model are well characterized by the derivation of mobile traffic distribution. Next, a nonlinear optimization framework for control traffic balancing is formulated to find the optimal forwarding paths from SD-switch and SD-BS to the controller in such a way the average control traffic delay in the whole network is minimized. Finally, the obtained delay results are further feedback to the placement decision, and trigger the adaptive control for better controller placement with the fulfillment of timely control message delivery.

(*2) Resource-Efficient Network Virtualization:* Network virtualization is essential to support infrastructure-as-a-service, thus enabling a wide range of emerging applications as mentioned above in Section 3.2. The virtual networks can be dedicated (i) to different network services/applications so that each service/application can be treated with customized and independent resource provisioning algorithms, (ii) to different network operators so that multiple operators can dynamically share the same network infrastructure along with the associated spectrum and infrastructure sharing, and (iii) to facilitate the cooperation and coexistence of different technologies, e.g., multi-radio RATs. Since wireless network resources are limited, resource-efficient wireless network virtulization solutions are highly desirable.

(*3) Distributed and Collaborative Traffic Classification:* The proposed SoftAir has the great capability to significantly improve the spectrum efficiency. Specifically, a distributed and collaborative traffic classifier has been proposed at SD-BS [31], which collaborates with a global traffic learner at the network controller to achieve fast, fine-grained and accurate traffic classification. The objective of this distributed traffic classifier is to identify the application, the QoS requirement, and the stochastic features associated with each traffic flow. With such traffic flow information, highly sophisticated and adaptive traffic engineering solutions can be adopted at both

Table 1
The comparison of existing W-SDN solutions.

W-SDN	SDN Architecture	Scalability	Network virtualization	SD Traffic engineering	Research community
OpenRoads [10]	SD-WiFi network	High	FlowVisor [14] & SNMPVisor	No specific solution	Academia/Industry
OpenRadio [11]	Programmable data plane	Low	No specific solution	No specific solution	Academia
CloudMAC [12]	SD-MAC in WANs	Low	- Dynamic spectrum use - On-demand AP - Downlink scheduling	Seamless AP switch-off system	Academia
Odin [13]	SD-MAC in WANs	Low	Hidden terminal mitigation	- Seamless mobility - Load balancing	Academia
CellSDN [9]	SD-CN	Low	Basic concept	No specific solution	Academia/Industry
ADRENALINE [15]	SD-CN	Low	OpenStack [21] (customer SDN controllers)	SD Optical OFDM system	Industry
SoftRAN [16]	SD-RAN	Low	No specific solution	Big-base station	Academia/Industry
Cloud-RAN [17]	SD-RAN	Low	Limited exploration	Collaborative PHY operations	Academia/Industry
SK Telecom [19]	SD-CN & SD-RAN	Low	Not supported	Collaborative PHY operations	Industry
DOCOMO [18]	SD-RAN	Low	Limited scheme, e.g., NOMA	Collaborative PHY operations	Industry
CONTENT [20]	SD-CN & SD-RAN	Low	Infrastructure manag. layer	Service orchestration layer	EU
SoftAir [3]	SD-CN & SD-RAN	High	- Network hypervisor - Wireless hypervisor - Switch hypervisor	- Collaborative & coordinate processing (PHY) - Collaborative scheduling (MAC) - SD mobility management (NW)	Academia/Industry

Table 2
Qualitative evaluation of NFV of existing W-SDN solutions.

W-SDN	NFV (Cloudification)	C/U-Plane decoupling
OpenRoads [10]	Simplified decoupling via OpenFlow	Simple extension to WiMAX AP
OpenRadio [11]	Enable multi-core DSP architectures	Focus on hardware design without control plane
CloudMAC [12]	VAPs for IEEE 802.11 stations	VAPs within virtual machines verse *dumb* WTPs
Odin [13]	LVAPs for IEEE 802.11 stations	No concrete solution for radio heads
CellSDN [9]	Abstract concept	Extension from SDN to W-SDN (Only CN)
ADRENALINE [15]	SDN integrated IT and network orchestrator (SINO)	Deploy end-to-end virtualized network functions
SoftRAN [16]	Rough NFV with big-base station abstraction	No concrete solution for radio heads
Cloud-RAN [17]	Rough NFV; suffer great I-Q transmissions	Fully & partial centralized architecture
SK Telecom [19]	Rough NFV; suffer great I-Q transmissions	Focus on hardware and software decoupling
DOCOMO [18]	Rough NFV; suffer great I-Q transmissions	Variant C-RAN; use decoupling in phantom cell
CONTENT [20]	Not much detailed NFV; suffer great I-Q transmissions	Infrastructure management layer: decouple across heterogeneous physical infrastructure
SoftAir [3]	Fine-grained NFV; solve redundant I-Q transmissions	Flexible platform for fully & partial centralized architecture (through software local control agents)

the BS and the network level. The proposed system consists of two components: (i) the local traffic classifiers in SD-BSs at the network edge and (ii) the global traffic learner at the network controller located at the CN. Therefore, SoftAir will take advantages of the centralized and computationally powerful network controller to jointly exploit the advantages of deep package inspection (DPI) [32] and semi-supervised machine learning [33].

Remark. The detailed comparison of state-of-the-art W-SDN solutions and SoftAir and the qualitative evaluation of NFV are summarized in Tables 1 and 2, respectively.

4. Conclusion

W-SDNs provide cellular networks with the needed flexibility to evolve and adapt according to the ever-changing network context for 5G cellular systems. An overview and qualitative evaluation of the state-of-the-art W-SDN solutions and the proposed SoftAir are deeply investigated with the highlighted key differences between these solutions. While the current researches have focused on enabling choice in different design axes, the question of how to make use of SDN and NFV concepts optimally for different scenarios still remains open.

Acknowledgment

This work was supported by the US National Science Foundation (NSF) under Grant no. 1547353.

References

[1] I.F. Akyildiz, A. Lee, P. Wang, M., Luo, W. Chou, A roadmap for traffic engineering in sdn-openflow networks, Comput. Netw. (Elsevier) J. 71 (2014) 1– 30.
[2] S. Jain, A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu, et al., B4: Experience with a globally-deployed software defined wan, Proceedings of the ACM SIGCOMM Conference, SIGCOMM'13, August 2013, pp. 3–14.
[3] I.F. Akyildiz, P. Wang, S.C. Lin, Softair: A software defined networking architecture for 5g wireless systems, Comput. Netw. 85 (2015) 1–18.

[4] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, J. Turner, Openflow: enabling innovation in campus networks, SIGCOMM Comput. Commun. Rev. 38 (2) (2008) 69–74.
[5] Common public radio interface (CPRI) specification v6.0; interface specification, 2013, Ericsson AB, Huawei Technologies Co. Ltd, NEC Corporation, Alcatel Lucent, and Nokia Siemens Networks GmbH & Co. KG. Std., http://www.cpri.info/downloads/CPRI_v_6_0_2013-08-30.pdf.
[6] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, J. Turner, Openflow: enabling innovation in campus networks, ACM SIGCOMM Comput. Commun. Rev. 38 (2) (April 2008) 69–74.
[7] M. Yu, J. Rexford, M.J. Freedman, J. Wang, Scalable flow-based networking with difane, in: Proceedings of the ACM SIGCOMM 2010 conference, ACM, New York, NY, USA, 2010, pp. 351–362.
[8] M. Casado, M.J. Freedman, J.P.J. Luo, N. McKeown, S. Shenker, Ethane: taking control of the enterprise, in: Proceedings of the ACM SIGCOMM 2007 conference, New York, NY, USA, 2007, pp. 1–12.
[9] L. Li, Z. Mao, J. Rexford, Toward software-defined cellular networks, in: Proceedings of the of European Workshop on Software Defined Networking (EWSDN), Darmstadt, Germany, 2010, pp. 7– 12.
[10] K.-K. Yap, R. Sherwood, M. Kobayashi, T.-Y. Huang, M. Chan, N. Handigol, N. McKeown, G. Parulkar, Blueprint for introducing innovation into wireless mobile networks, in: In VISA 10: Proceedings of the second ACM SIGCOMM workshop on Virtualized infrastructure systems and architectures, New York, NY, 2010, pp. 25–32.
[11] M. Bansal, J. Mehlman, S. Katti, P. Levis, Openradio: a programmable wireless dataplane, in: Proceedings of the of the first workshop on Hot topics in software defined networks (HotSDN '12), Helsinki, Finland, 2012, pp. 109–114.
[12] P. Dely, J. Vestin, A. Kassler, N. Bayer, H. Einsiedler, C. Peylo, Cloudmac-an openflow based architecture for 802.11 mac layer processing in the cloud, in: Proceedings of the Globecom Workshops (GC Wkshps), 2012 IEEE, 2012, pp. 186–191.
[13] L. Suresh, J. Schulz-Zander, R. Merz, A. Feldmann, T. Vazao, Towards programmable enterprise wlans with odin, in: Proceedings of the First Workshop on Hot Topics in Software Defined Networks, in: HotSDN '12, 2012, pp. 115–120.
[14] R. Sherwood, G. Gibby, K.K. Yapy, G. Appenzellery, M. Casado, N. McKeowny, G. Parulkar, FlowVisor: A Network Virtualization Layer, Technical Report, 2009.
[15] Centre Tecnologic Telecomunicacions Catalunya (CTTC), SDN/NFV Cloud Computing Platform and Core Network for 5G Services, Technical Report, http://networks.cttc.es/ons/adrenaline/sdnnfv-cloud-computing-platform-and-core-network-for-5g-services/
[16] A. Gudipati, D. Perry, L. Li, S. Katti, Softran: software defined radio access network, in: Proceedings of the first workshop on Hot topics in software defined networks (HotSDN '13), Hong Kong, 2013, pp. 25–30.
[17] China Mobile Research Institute, C-RAN White Paper: The Road Towards Green RAN, Technical Report, 2014.
[18] NTT DOCOMO, INC, DOCOMO 5G White Paper: 5G Radio Access: Requirements, Concept and Technologies, Technical Report, 2014.
[19] SK Telecom, SK Telecom 5G White Paper: SK Telecom's View on 5G Vision, Architecture, Technology, and Spectrum, Technical Report, 2014.
[20] 2012-2015, (Project CONTENT FP).
[21] OpenStack: The Open Source Cloud Operating System. [Online]. Available: https://www.openstack.org/software/.
[22] J. Wu, Z. Zhang, Y. Hong, Y. Wen, Cloud radio access network (c-ran): a primer, IEEE Netw. 29 (1) (2015) 35–41.
[23] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra, S. Rangarajan, Radio access network virtualization for future mobile carrier networks, IEEE Commun. Mag. 51 (7) (2013) 27–35.
[24] C.-Y. Hong, S. Kandula, R. Mahajan, M. Zhang, V. Gill, M. Nanduri, R. Wattenhofer, Achieving high utilization with software-driven wan, Proceedings of the ACM SIGCOMM 2013, August 2013, pp. 15–26.
[25] Huawei technologies sdn showcase at sdn and openflow world congres2013, http://www.eantc.de/fileadmin/eantc/downloads/events/2011-2015/SDNOF2013/EANTC-Huawei_SDN_Showcase-White_Paper_Final_Secure.pdf.
[26] S.-C. Lin, P. Wang, M. Luo, Control traffic balancing in software defined networks, Comput. Netw. (2015).
[27] B. Pfaff, J. Pettit, K. Amidon, M. Casado, T. Koponen, S. Shenker, Extending networking into the virtualization layer, in: ACM Workshop on Hot Topics in Networks 2009, 2009.
[28] Active RAN Sharing Could Save 60 Billion for Operators. [Online]. Available: http://www.cellular-news.com/story/36831.php.
[29] A.X. Porxas, S.-C. Lin, M. Luo, Qos-aware virtualization-enabled routing in software-defined networks, in: IEEE ICC 2015, London, UK, 2015.
[30] H. Jin, D. Pan, J. Liu, N. Pissinou, Openflow-based flow-level bandwidth provisioning for cicq switches, IEEE Trans. Comput. 62 (9) (2013) 1799–1812.
[31] Q. Hu, P. Wang, I.F. Akyildiz, M. Luo, A framework for qos-aware traffic classification using semisupervised machine learning in sdns, 2015. Under preparation.
[32] A.W. Moore, K. Papagiannaki, Toward the accurate identification of network applications, in: Passive and Active Measurement (PAM05, 2005.
[33] B.F. Lo, I.F. Akyildiz, Reinforcement learning for cooperative sensing gain in cognitive radio ad hoc networks, ACM-Springer Journal of Wireless Networks 19 (6) (2013) 1237–1250.

Ian F. Akyildiz (M'86-SM'89-F'96) received the B.S., M.S., and Ph.D. degrees in Computer Engineering from the University of Erlangen-Nurnberg, Germany, in 1978, 1981 and 1984, respectively. Currently, he is the Ken Byers Chair Professor in Telecommunications with the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, GA USA; the Director of the Broadband Wireless Networking (BWN) Laboratory and the Chair of the Telecommunication Group at Georgia Tech. Since 2013, he is a FiDiPro Professor (Finland Distinguished Professor Program (FiDiPro) supported by the Academy of Science) with the Department of Electronics and Communications Engineering, at Tampere University of Technology, Finland, and the founding director of NCC (Nano Communications Center). Since 2008, he is also an honorary professor with the School of Electrical Engineering at UniversitatPolitcnica de Catalunya (UPC) in Barcelona, Catalunya, Spain, and the founding director of N3Cat (NaNoNetworking Center in Catalunya). Since 2011, he is a Consulting Chair Professor at the Department of Information Technology, King Abdulaziz University (KAU) in Jeddah, Saudi Arabia. He is the Editor-in-Chief of Computer Networks (Elsevier) Journal, and the founding Editor-in-Chief of the Ad Hoc Networks (Elsevier) Journal, the Physical Communication (Elsevier) Journal and the Nano Communication Networks (Elsevier) Journal. He is an IEEE Fellow (1996) and an ACM Fellow (1997). He received numerous awards from IEEE and ACM. His h-index is 88 and the total number of citations is above 69K due to Google scholar as of March 2015. His current research interests are in wireless sensor networks in challenged environments, 5G Cellular systems, nanonetworks, Terahertz Band, and software defined networks.

Shih-Chun Lin (S'08) received the B.S. degree in electrical engineering and the M.S. degree in communication engineering from National Taiwan University in 2008 and 2010, respectively. He is a graduate research assistant in the Broadband Wireless Networking Laboratory (BWN Lab), School of Electrical and Computer Engineering, Georgia Institute of Technology. Currently, he is working toward the PhD degree in electrical and computer engineering under the supervision of Prof. Ian F. Akyildiz. His research interests include wireless underground sensor networks, software defined networking, large machine-to-machine communication, cognitive radio networks, and statistical scheduling in wireless systems.

Pu Wang (M'05) received the B.S. degree in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 2003, the M.Eng. degree in computer engineering from the Memorial University of Newfoundland, St. Johns, NL, Canada, in 2008, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, in 2013. He is currently an Assistant Professor with the Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS. Dr. Wang received the BWN Lab Researcher of the Year award in 2012, Georgia Institute of Technology. He received the TPC top ranked paper award of IEEE DySPAN 2011. He was also named Fellow of the School of Graduate Studies, 2008, Memorial University of Newfoundland. His research interests include wireless sensor networks, cognitive radio networks, software defined networks, nanonetworks, multimedia communications, wireless communications in challenged environment, and cyber-physical systems.