Review article

IoT survey: An SDN and fog computing perspective

Ola Salman*, Imad Elhaj, Ali Chehab, Ayman Kayssi

Department Electrical and Computer Engineering, American University of Beirut, Beirut 1107 2020, Lebanon

1. Introduction

One of the all-time most impactful innovations is the Internet. Internet has permitted the interconnection of all traditional computing devices and it was natural for this desire for access and control to extend to non-traditional devices. Here came the evolution into Internet of Things (IoT). Mentioned seventeen years ago by Kevin Ashton [1], IoT draws the lines of the second digital revolution [2,3]. Cisco expected that, by 2020, 50 billion objects would be connected to the Internet [4]. This large scale is one of the unavoidable challenges for the IoT domain. The high scalability is accompanied with an increased complexity in the management of this large number of things/gateways, and network devices. Managing all these devices in the traditional way (manually and each device separately) is no longer viable.

As the Metcalfe’s law states, the importance of a communication network increases exponentially with the number of connected devices [5]. Therefore, with billions of connected things in the future network, the IoT value is extremely high [6]. In addition, IoT is depicted as one of the most disruptive technologies [7,8]. Many firms and technology leaders (Intel, Microsoft, Cisco, InterDigital, etc.) have taken note of the IoT economical value [9], and put serious efforts to enable IoT real deployment (Table 1 lists some of the important ongoing projects). However, this drive to develop IoT solutions has resulted in proposing disjoint ones. Lacking interoperability between the different IoT platforms limits their potential. We all know that the root enabler of the Internet success and wide adoption is its openness and its standardized architecture. Having different IoT architectures and platform resulted in having heterogeneous silos of networks. In addition to this kind of heterogeneity, different formats of data are used, and different types of communication technologies are invoked. This makes the IoT a vertically fragmented network. Therefore, the heterogeneity is another important challenge facing IoT.

Moreover, the large number of connected devices will naturally result in enormous amount of data, which challenges the ability of today’s networks to handle. The current centralized paradigm of data processing and storage is not feasible. New ways to analyze filter and aggregate this data at the network edges will be essential in any upcoming IoT solution. The IoT "Big Data" is not only about the size of the generated data, but it is more about the variety of this data in terms of type, semantic, frequency, place and time.

Finally, security and privacy guarantees present one of the most important challenges that effectively hinders any real IoT wide deployment. In addition to the current security vulnerabilities, IoT poses new ones.

In the light of the cited challenges, there is a need for new approach to networking. Software Defined Networking (SDN), a new networking paradigm, aims to separate the control and data planes. This separation provides the network controller with a global view of the network, facilitating traffic engineering and network management at runtime [10]. On the other hand, fog comput-
ing (a cloud computing complement) aims at bringing the cloud to the network edge making it more scalable and more responsive. In this survey, we will investigate how these technologies have been applied in the IoT field and how their application will enable the IoT wide deployment.

In the literature, several surveys have tackled different IoT related subjects (see Table 2): IoT applications, challenges, and opportunities [21–31], IoT frameworks [32–37], IoT security [38–43], IoT standardization [44–48], SDN application in IoT [49–54], IoT and cloud integration [55,56].

However, the existing surveys do not comprehensively review the main IoT challenges. The Internet already presents QoS and security related challenges, but in the IoT case, some of the existing challenges become more crucial. Thus, the existing work includes specific and non-specific IoT challenges. In addition, the IoT challenges are listed without including the related proposed solutions. However, in this work, we presented the four main IoT-specific challenges and we reviewed the proposed solutions coping with these challenges. In this context, new technologies emerge in the network, communication, and IT domains. These technologies can enable innovative IoT applications and help in coping with many of the IoT challenges. However, the existing work does not cover the most recent technologies and the role of these technologies in alleviating the IoT challenges. Thus, this work presents the most recent enabling technologies, and how these technologies can be employed to cope with the presented IoT challenges. Specifically, this paper reviews the application of SDN, NFV, cloud computing, and fog computing to handle the main IoT challenges.

This paper is organized as follows: In Section 2, we discuss the most relevant IoT definitions and we list the main related IoT concepts. In Section 3, we investigate the IoT enabling technologies. In Section 4, we review the most important IoT, SDN, NFV, and edge computing standardization efforts. In the subsequent sections, we present the IoT main challenges (as shown in Table 3): IoT security, IoT Big Data, IoT heterogeneity, and IoT scalability, and the corresponding SDN/NFV and cloud/edge-based solutions. Thus, in Section 5, we review the IoT security related work and we show how SDN can alleviate the IoT security concerns. Section 6 reviews the IoT “Big Data” and the application of cloud and fog computing to manage it. The IoT gateway is an essential part to cope with the heterogeneity challenge, so in Section 7, we review the propositions of IoT gateways. IoT scalability imposes new architectural considerations, so Section 8 reviews the most known IoT architectures and the SDN integration into a general IoT architecture. In Section 9, we present the main limitations of the current IoT solutions and we make some recommendations for the future research directions. Finally, we conclude in Section 10.

2. IoT definition

Beyond the IoT hype, a real definition is essential to highlight the characteristics of this new concept [57]. Several definitions have been proposed resulting in a storm of terms and definitions. Important work is being done by the IEEE Internet initiative in order to find a conceptual IoT definition [58]. ITU defines IoT as being an infrastructure that will connect physical and virtual devices [59]. IETF defines IoT as being the Internet that considers TCP/IP and Non-TCP/IP suites at the same time and the things as being “objects” identified by unique addresses [60]. IEEE, in its special report on Internet of Things, defines it as a network that connects devices having sensing capabilities [58]. In [58], the IEEE Internet initiative gives its own definition as follows: The Internet of Things is a network that connects uniquely identifiable virtual and physical devices, using existing or new communications protocols. These Things are dynamically configurable and have interfaces that must be accessible distantly through the Internet [58].

Actually, IoT is not the exclusive name for this new concept. In the IoT storm, there are different and confusing terms such as: Machine-to-Machine (M2M), Industrial Internet of Things (IIoT), Internet of Anything (IoA), Internet of Everything (IoE), Web of Things (WoT), and Social Internet of Things (SIoT). An inclusion relation can be established between these different concepts as shown in Fig. 1. In the following, we will try to present the definitions of these concepts IoE, coined by Cisco [61], consists of connecting things, devices, humans, and data to a global network [62]. IoA, presented in [63], consists of connecting not only the existing well-known things as implied in the IoE definition, but also it refers to connect all possibly “imagined” things. M2M is a subset of IoT [64], which includes the M2M communications as well as Machine-to-Human interaction. ETSI defines M2M in [65] as: an automated communication between two devices without a human intervention [65]. While IoT focuses on the physical objects’ representation, M2M is connectivity centric. Thus, moving from M2M to IoT necessitates further considerations [66].

While IoT applications tackle different human life domains, the industrial field remains one of the most critical ones. Applying IoT in the industrial domain requires careful attention and special efforts [67,68]. Talking about industries means that we include businesses ranging from small ones to large ones. Security and privacy are the most challenging issues in this context [69]. Recently, a consortium for the industrial IoT, IIoT, has been established [70]. This consortium, founded by AT&T, Cisco, GE, IBM, and Intel in March 2014, aims at pushing the standardization in this area [71].

In the Auto-ID Labs white paper presented in [72], a comparison between IoT and Web of things, WoT, is performed. It is stated that IoT is a wider concept than WoT, having structural concerns (e.g. unique identity for the things), which cannot be resolved by the web technology. Essentially, WoT is a web framework to which
Table 2
IoT surveys summary.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Reference</th>
<th>Contributions</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>IoT applications, challenges, and opportunities</td>
<td>[21–31]</td>
<td>The main IoT benefits, applications (smart home, healthcare, connected cars), and challenges are presented.</td>
<td>The challenges are just mentioned without presenting the enabling technologies and solutions.</td>
</tr>
<tr>
<td>IoT frameworks</td>
<td>[32–37]</td>
<td>The different proposed IoT frameworks are presented.</td>
<td>The SDN and fog-based frameworks are not included.</td>
</tr>
<tr>
<td>IoT Security</td>
<td>[38–43]</td>
<td>The IoT security challenges and the proposed protocols are presented.</td>
<td>The SDN benefits in terms of security are not considered.</td>
</tr>
<tr>
<td>IoT Standardization</td>
<td>[44–48]</td>
<td>The standardization efforts in the IoT domain are reviewed.</td>
<td>The standardization efforts for the new emerging technologies like SDN and fog computing are not presented.</td>
</tr>
<tr>
<td>SDN application in IoT</td>
<td>[49–54]</td>
<td>The SDN application at different IoT levels are presented.</td>
<td>The role of SDN in alleviating the IoT challenges is not included.</td>
</tr>
<tr>
<td>Cloud/Edge computing for IoT</td>
<td>[55,56]</td>
<td>The cloud related application to enable different IoT applications are presented.</td>
<td>The focus is on the big data related challenge. However, the networking aspect of the different data nodes is not considered (the application of SDN for data networking).</td>
</tr>
</tbody>
</table>

Table 3
IoT challenges, benefits, and limitations.

<table>
<thead>
<tr>
<th>IoT challenges</th>
<th>Benefits</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability</td>
<td>• Connecting new kinds of devices</td>
<td>• Management complexity</td>
</tr>
<tr>
<td></td>
<td>• Gain more control over the connected devices</td>
<td>• Network capacity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• QoS</td>
</tr>
<tr>
<td>Big Data</td>
<td>• Enabling innovative applications</td>
<td>• Big Data management</td>
</tr>
<tr>
<td></td>
<td>• Getting useful insights from Big Data analysis</td>
<td>• Data centralization: high latency, redundancy, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data at the network edge: security, management, networking, etc.</td>
</tr>
<tr>
<td>Heterogeneity</td>
<td>• Integrating different IoT vertical silos</td>
<td>• Interoperability</td>
</tr>
<tr>
<td></td>
<td>• Integrating different communication technologies, devices’ types, data types, etc.</td>
<td></td>
</tr>
<tr>
<td>Security and Privacy</td>
<td>• Enabling innovative applications using sensitive data</td>
<td>• New types of attacks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Private data inspection</td>
</tr>
</tbody>
</table>

the things are connected through the Internet and have their collected data pushed to it. Web data analysis, and user interfaces are keys to provide services that enable innovative applications. The web can be used to access data but the communication between devices, automation, auto-configuration, and management capabilities are outside the scope of the existing web.

**IoT**, which allows the things to have their social networks [73,74], is a related concept to WoT. Inheriting the success of social networks (e.g. Facebook), that can be considered as “banks of data”, the socialization concept can be employed in the IoT context. The projection of the IoT world in the social one results in the projection of the things into the social world, which requires new things’ definitions (social objects) [75].

The Future Internet (FI) is a global network that will encompass all the above-mentioned networks. Six principles (C6) will enable this innovation. The C6 annotation refers to: Connectivity, Content, Cloud, Context, Collaboration, and Cognition [76]. In such a network, mobile and constrained things will be connected to the Internet generating huge amount of data. This data, handled by advanced cloud-based technologies, will shift context-aware behavior into a collaborative environment between the different things. The analysis of this data will result in a cognitive world. Effectively, IoT is an essential part of the Future Internet [77]. Accordingly, it is important to show the distinction between these concepts in the aim to solve their specific problems and to allow their realization in the near future [78].

3. **Enabling technologies**

In this section, we present the recent technologies designated to play an essential role in the IoT realization.

3.1. **SDN & NFV**

SDN, an emerging technology in the network domain, aims at separating control and data planes. The control plane consists of the SDN controller (Network Operating System (NOS)) which has the role of network orchestration; most of the computations are done there, which gives it a special importance being the network brain. The data plane consists of the network devices (routers/switches) being responsible for simple matching operations to know how to forward the packets. These simple devices forward to the controller every packet they do not know how to act upon.

SDN is not the first attempt towards separating data forwarding and network strategical computation, and it is not the first trial to softenize the network functions. The history of programmable networks dates back to the early 90s. Several attempts to apply programmability and automation in the network domain have been conceived (e.g. ATM) [79]. However, SDN is the most promising one. OpenFlow, the first standardized southbound interface, has presented a primary insight onto the network programmability ef-
The SDN architecture, as presented in the recent IRTF RFC 7426 drafted by the SDN Research Group (SDNRG) [80], consists of four layers: data layer, control layer, management layer, and application layer as shown in Fig. 2. This architecture provides network flexibility, dynamicity, and management capabilities. In this context, the question about how SDN can enable IoT arises.

Therefore, one can argue that it is a bad idea to propose a centralized architecture for this highly scalable network, or that applying the SDN paradigm to it has a retro effect that returns us back to the era where the centralization paradigm was pioneering the telecom domain. However, these assertions are not very accurate. SDN centralized control cannot be compared to a central telecom switch. The SDN centralization is a logical concept more than being a physical one. In this context, the distributed control scheme came to defend this argument [81-82]. Several controllers have enabled control distributiveness into their architecture: Disco [83], Onix [84], ONOS [85], and OpenDaylight [86], etc. In this case, the East/Westbound [87] interfaces are responsible for connecting the distributed controllers’ instances. Besides, the southbound interfaces provide the control over the network. OpenFlow is the most known southbound interface. The OpenFlow operational flow is summarized in Fig. 3.

Upon receiving a packet, the switch performs filter matching on the header fields. If there is a corresponding entry in this table, it takes action based on the entry’s action part. Otherwise, it forwards the packet to the controller. When it receives the packet from the switch (PacketIn), the controller takes the forwarding decision and downloads the corresponding rule to the switch (PacketOut). The controller basic modules are: topology manager, link manager, decision making, data control, and data storage [88]. The controller learns about the network links (between switches and switches and hosts) using the Link Layer Discovery Protocol (LLDP). The link discovery module provides this information to the topology manager module that is responsible of constructing/updating the network topology database.

On top of the control layer, the application layer resides. The communication between the control and application layers is performed through the northbound interfaces. These interfaces, which give the application access to the network collected data, offer most of the SDN benefits.

The IoT realization is mostly hindered by the Internet management complexity issue preventing the dynamic deployment of new services. The control is fully distributed; so, reconfiguring the network and adding new features will be exhaustive if done in the traditional way. Using SDN, this task becomes much simpler; the control centralization provides the controller with a global view of the network, giving it the power to hide the management complexities and to have more control over the network. QoS guarantee, heterogeneity, security and privacy concerns, communication resilience, and big data management are tasks that can be alleviated by the SDN introduction [89-90].

On the other hand, after revealing its innovative value in the IoT domain, the virtualization finds its way into the network domain with the SDN proliferation. NFV and SDN are complementary technologies. SDN with NFV (or SDNv2 [91]) allows the virtualization of the network functions in a way similar to what we have seen in the computing domain; the same network infrastructure can be used by different applications. The network is divided into slices in this case, and each slice has to support certain flow. This allows a fine-grained services categorization and offers a security enhancement solution.

Additionally, NFV plays an important role in the IoT domain [92]. Coupled with SDN, this technology has the capability to handle the IoT requirements in terms of QoS guarantee, traffic engineering, defeating heterogeneity, and providing security services. Additionally, NFV helps to cope with the IoT high scalability challenge [93]. Due to the limited network capacity, the increase in the number of connected devices poses network constraints that cannot be met, especially at peak time load. However, upgrading the existing network infrastructure to support higher capacity is expensive in terms of both OpEx and CapEx. In this context, virtualization provides elasticity that helps in an optimized use of the limited hardware resources at low network load permitting the sharing of the network infrastructure between different service providers and different network services/functions. Additionally, NFV allows to borrow network resources as needed at runtime preventing the waste of resources if designed to handle the peak load [94].

### 3.2. Cloud computing (X-as a Service)

Cloud computing has permitted advancement in the network and telecommunication domains. Relying on the “pay-as-you-go” paradigm, it enables the reduction of both OpEx and CapEx. In the IT domain, several big companies have built their own cloud systems (e.g. Microsoft [95], Google [96], Apple [97], etc.), and some have exploited it from an economical perspective. Cloud was a revolution in the IT domain and it draws a new line in the telecommunication domain. The idea behind the cloud is to borrow computing facilities from the cloud and to pay as you use the pro-
vided services. Three main services were provided by the cloud: SaaS (Software as a Service), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service).

In the IoT case, the limitations in terms of processing and computing capabilities are more pronounced. Employing cloud computing, in the IoT domain, is beneficial having limited storage and computational power devices. These two revolutionary technologies (cloud and IoT) are fundamental and complementary for the future network era [98]. Different IoT functions have been proposed to be part of the cloud services such as: sensing and actuating as a service [99], sensing cloud [100], sensing as a service [101], building the environment for the Internet of Things as a service (BETaaS) [102], etc.

Another complementary aspect of cloud and IoT is the fact that IoT calls for computation, storage, and communication resources remedy due to the constrained nature of things. On the other hand, the cloud providers need new market domains [103]. Combining IoT and cloud computing brings benefits as well as challenges. Two angles of convergence are presented in [104]: cloud-based and IoT-based. While most of the work discusses cloud-based IoT services, the idea of IoT-centric approach was new. The IoT-centric cloud aims to push cloud functionalities to IoT network edge. This is similar to what is called "fog computing". However, the cloud-centric model presents some challenges regarding the cost of transporting the data through the core network, the high latency, and the single point of failure (reliability). Thus, a new way to manage the data is needed while keeping the cloud as a backend.

3.3. Fog computing/MEC/Cloudlet

Rapid mobility patterns, high throughput, reliable sensing, reliable control and actuation, very low latency, big data management, different levels of real-time analytics, and data aggregation are main IoT requirements that cannot be met concurrently by the cloud technology [105]. Low latency, low jitter, mobility support, location awareness, augmented reality, geo-distributiveness, and multi-tenancy applications support (IaaS) are common characteristics provided by edge computing in its different flavors (fog, MEC, and cloudlet) as shown in Fig. 4.

In this context, edge cloud computing was proposed to push the data collection, processing, and analysis to the network edge [106,107]. These edges will not be very powerful nodes; they would just complement the cloud. Essentially, edge computing is about to know which data has to be analyzed at which point (i.e. which data has to stay at the edge and which data has to be pushed to the cloud) [107–109]. Thus, the collaboration between cloud and edge is mandatory [110].

In the shadow of edge computing, comparable new trends aroused (Fig. 5) [111,112]. Fog, cloudlet, and Multi-access Edge Computing (MEC) are three significantly related concepts. In the following, we will try to investigate the subtle differences between these technologies. In this context, the work done in [113] was a good reference for such a comparison.

3.3.1. Fog

Inspired by the natural phenomenon of having fog and clouds where fog are closer to the ground [113], fog computing is meant to be the cloud at the network edge in the IoT networks [114–116].

Fog computing, coined by Cisco in 2012, is an extension of the cloud to the network edge [117]. The fog related characteristics such as low latency, geo-distribution, location awareness, support for mobility, support for ubiquitous access, and support for heterogeneity, present basic requirements for a wide range of IoT services and applications [118]. Additionally, the IoT high scalability imposes federated network management and thus call for new network and data technologies to enable IoT data processing at the network edge [119–121].

In [122], the authors shed light on the relation between fog and cloud. Fog and cloud are complementary technologies and none of them replaces the other. The differentiation between fog and cloud is meant to be in the type of required data and the speed with which data must be processed. Local information can be served by fog nodes and global information can be served by the cloud. The short distance to end users makes the fog distributed platform more suitable for IoT applications while the cloud is relatively farther away [123].

“Why Fog and Why Now?” delay, cognition, agility, and efficiency are defined to be the main reasons for the fog invocation [122]. Additionally, reliability, fault tolerance, and privacy are presented to be fog related benefits [124]. Essentially, cognitive assistance gives rise to distinctive services to be provided by IoT. Cognitive assistance will be the "killer app" for mobile computing in the next decade. However, human perception is sensitive to latency. To gain user satisfaction, such applications have to benefit from fog computing emergence to deliver low latency and high-performance processing [125]. Additionally, other innovative applications (e.g. healthcare, smart cars, etc.) will benefit from the fog computing emergence [126,127].

However, the fog nodes distributed pattern imposes new networking issues [128,129]. While the data is centralized at one point in the cloud case, the data nodes are distributed in the fog case. Inducing ways to optimize data correlation between these nodes is critical. Other issues are also encountered in the fog domain such as security and privacy, provisioning and resource management, offloading, charging and accounting, and QoS guarantees in terms of capacity, storage, bandwidth, connectivity, and reliability [130–133].

3.3.2. MEC

An ETSI white paper introduced its Industry Specification Group (ISG) intended to create MEC specifications [134]. This initiative aims at merging the IT and telecom domains to provide cloud-based services at the mobile network edge. In [134], sketches for possible MEC scenarios are presented after listing their benefits and use cases. ETSI is involved in many initiatives such as: NFV, 3GPP, OneM2M, etc. So, it has the incentive to relate new technologies to old ones taking advantage of mature old techniques and providing backward compatibility. In this context, MEC uses cases have been applied to the 3GPP mobile architecture. More recently, MEC solutions have been proposed in the 5G domain [135–137]. Consequently, MEC is one the IoT enablers [138].
MEC consists of leveraging edge nodes to enhance mobile devices capabilities. The developers of mobile applications have to account for screen sizes, memory capacity and processing power of different devices, and they must consider the different situations when the processing is done locally or remotely. Handling connection (mobility management) is another issue to deal with. Additionally, while content providers can benefit from giving the MEC’s providers insights about users’ preferences, it is critical for application developers to protect users’ anonymity and privacy. The data synchronization and the mutual trust are the main challenges in this case [139]. Additionally, computation offloading is one of the main MEC challenges having processing and memory limitations [140–143].

3.3.3. Cloudlet

Known also as follow me cloud [144], mobile micro-cloud [145], and mobile cloud computing (MCC), cloudlet, coined by a research group at Carnegie Mellon University, is defined as a small datacenter at the Internet edge [146].

In [147], the mobile edge clouds or cloudlets are considered collocated with the base stations. The authors search to optimize the service migration decision. Their approach is based on the distance between user and base station for simplicity. The optimization problem is formulated using the Markov Decision Process (MDP).

Thus, the main differences between cloudlet and cloud are the rapid provisioning, the fast hand-off, and the cloudlet discovery [148]. Additionally, cloudlets are dispersed at the network edges while the cloud servers are centralized at the core. The cloudlets are managed in an autonomic way while the cloud is managed in an administrative central way. The local edge nodes serve a few number of users giving them augmented reality experience while the cloud case connects a huge number of users [146]. While Cloudlet and MEC were conceived in the mobile domain, the fog computing term is used in the IoT domain.

While cloud computing and edge computing are mainly intended for IoT data processing, storage, and management, there is a need for new networking paradigm to manage the interconnection between the datacenters and/or edge nodes. SDN initially was applied mainly within the datacenters. SDN provided agility and flexibility in deploying and managing the needed network resources for VMs allocation. The need for SDN is more pronounced in the case of distributed cloud nodes (edge computing). In this case, the interconnection and management of the distributed edge nodes call for the SDN manageability. Thus, SDN, NFV, and edge computing can together be employed to handle both IoT data and networks.

3.4. Cellular IoT (5G)

The communication era has witnessed a distinctive evolution; from networks handling analog voice services to fully IP-enabled mobile networks. The first mobile generation (1G) was a revolution supporting user mobility after the fixed telecom network. New services have emerged with newer versions. With the second digital mobile generation (2G), new services and applications have emerged (i.e. text messaging). However, with the third and fourth generations (3G and 4G), distinctive applications have appeared (apple Siri, google glass, etc.) paving the way towards new mobile telecommunication epoch. Cell phones will dominate the future Internet [149]. “Horizon 2020 and beyond” is the tag of the upcoming new era in the telecommunication domain. 5G is not depicted to be an evolution of the previous mobile network generations (2G/3G/4G/LTE). It is more of a revolution that will change our way of life. 5G is supposed to be the second industrial revolution. Very low latency, high throughput, reliability, security, and high mobility are the characteristics of this upcoming technology [150]. This revolution will enable the cellular IoT paradigm [151]. Main IoT requirements are to be met by this new mobile network. IoT is expected to be integrated in the 5G mobile network [152]. The no cell communication pattern will be supported encompassing the Device-to-Device (D2D) direct communication reducing the signaling and connection time. There is an effort to integrate Machine Type Communication (MTC) into the 3GPP mobile network architecture supporting essentially Human-to-Human (H2H) communication. The mobile network capacity to handle mobile wireless communication makes it suitable for the new emerged communication type. However, the high number of connected devices...
will need a new management paradigm. Overhead in terms of signaling and communicated data must be considered in any mobile network bearing M2M based services.

With the proliferation of smart mobile phones and the dense access to the access nodes, a new paradigm for managing radio access network is needed. The old way adopted in mobile network relies on distributed RAN management. However, the distributiveness has its limitation with high density, high scalability, and low latency requirements. Migrating the control to a centralized entity with a global view of multiple cells will handle better the handover management, power allocation, and interference management. Nevertheless, the decision on downlink allocation per resource block can be given to the local RAN control unit [153]. In this context, applying SDN to the telecom domain has been considered as shown in [154]. Table 4 summarizes the most relevant work done to apply SDN in the mobile network domain.

In [155], there is a proposal for a generalized architecture for mobile networks integrating SDN & NFV. Applying SDN in mobile networks is constrained by the ability of this technology to provide a clear path of migration (support co-existence of different generations), security, QoS monitoring, service provisioning, and cost reduction. The main functions to be softwareized are mainly the mobile network control functions, i.e., MME, HSS, PCRF, and S/P-GW. Additional functions include transport, load balancing, security, policy, charging, monitoring, and QoE or resource optimization.

Recently, a 5G operating system (OS) consisting of three levels of control, device controller, edge controller, and orchestrator controller, was proposed in [150]. The device controller encapsulates certain level of intelligence at the device level (machine learning). However, due to the constrained power conditions, this device might call for higher level of control to optimize its power resources. The edge controller is responsible for L2–L3 functions (forwarding routing, QoS provisioning, mobility management, and charging functions). The orchestrator is the highest level of control, it has the role of managing the cloud resources (links, storage, and memory), the allocation, and the provisioning of the VMs. Everything as a service (XaaS) and IoT are the main profiting technologies from the 5G OS conception.

Consequently, new mobile network architecture (5G) with emergence of new technologies such as SDN & NFV and cloud/fog computing will have a major impact on enabling IoT.

3.5. WSN

Wireless Sensor Network (WSN) is one of the utmost IoT application domains. Being able to wirelessly collect data from sensors spread and integrated into different things is an essential IoT requirement. Also having control over these things thanks to embedded actuators is as important. So, WSN is an essential IoT enabler. However, the wireless technology presents many challenges in terms of security due to the ubiquitous network access and in terms of QoS guarantee due to the unpredictable number of connections (mobility) and the environmental influence (interference).

Essentially, routing is one of the obstacles that was tackled widely in the WSN domain. However, conceiving a routing protocol that is energy efficient, supporting load balancing, and dynamically adaptable to network changes, is not a simple task in the distributed traditional way. Centralizing the control at the master and center node levels will be beneficial in terms of both robustness and energy efficiency [162].

From this perspective, SDN has been essentially applied to the wired networks. Indeed, there are many attempts to apply it to the wireless networks [163]. Sensor OpenFlow [164], and SDWN [165,166] are examples of applying OpenFlow/SDN to the WSN domain. The main idea is to enable SDN on the access points through Open vSwitch along with a centralized SDN controller [167]. This controller has the role of managing routing, flow scheduling, and interference management. Additionally, security related functions are assigned to this central controller.

4. Standardization

4.1. IoT standardization

Standardization is key to achieve any new technology’s wide adoption. Having disjoint platforms, architectures and protocols undermine their utility. A standardized IoT architecture is key for the IoT wide deployment in addition to the standardization at the communication level [168]. The TCP/IP standard was the enabler of the Internet revolution. Revisiting its architecture, we found that most of its protocols at different layers are not designed for the IoT case. The “things” in IoT might be constrained devices, so the power consumption at different layers must be taken into consideration [169]. Additionally, the IP protocol itself is overwhelmed by the big number of connected things (already IPv4 addresses pool has been exhausted). In addition, the security related protocols, already not widely adopted in the current Internet network due to their expensive cost (overhead), must be revisited as well.

Under the IETF guidance, several working groups (WGs) have been established in the aim of standardizing new IoT protocols or adapting the existing TCP/IP protocol stack to be suitable for IoT (Fig. 6). CoAP, a lightweight HTTP version and CoRE, which is based on the REST web technology for constrained IoT devices, are application layer protocols. DTLS is a security transport layer protocol suitable for constrained devices (running over UDP). Under the RoLL (Routing over Low power and Lossy networks) WG, the IPv6 Routing protocol for Low power and Lossy Networks (RPL) has been developed. In a parallel effort, IEEE has developed the IEEE
802.15.4, which covers the physical and MAC layers of the TCP/IP stack. The Bluetooth Low Energy (BLE) is another radio communication standard based on IEEE 802.15.1; it is characterized by its low energy and fair data rate, which makes it suitable for some of the IoT applications [170]. The ZigBee alliance builds on top of the IEEE 802.15.4 its own architecture for low power communication network.

However, integrating these protocols in the current Internet infrastructure will not help in overcoming the cited IoT challenges. Perhaps, they will add complexity in terms of management and protocols interoperability. Furthermore, The IoT challenges need to be considered profoundly and correspondent solutions must be engineered harmoniously in one IoT architecture.

4.2. SDN and NFV standardization

The standardization efforts in the SDN and NFV domains are not conducted by a single entity. However, many Standards-Developing Organizations (SDOs), industry consortium, and open development organizations have participated in developing SDN and NFV related standards. The Internet Society (ISOC) has two working groups: IETF and IRTF, that are working on SDN related standards. Interface to routing systems (I2RS) and service function chaining are two groups under the IRTF organization that are working on SDN related specifications. Additionally, IRTF has published an RFC titled “Software-Defined Networking (SDN): Layers and Architecture Terminology” (RFC 7426, January 2015) [80]. The ITU-T has four groups (SG11, SG13, SG15, SG16) working on SDN related projects. Additionally, ETSI was the leader in proposing an NFV standardized architecture. Other open development initiatives like OpenDaylight (under the Linux Foundation directory) and Open Platform for NFV (OPNFV) are working on open standards and open source projects that are designed to play an essential role in the business domain [171].

4.3. Edge computing standardization

The Multi-access Edge Computing (MEC) initiative is an Industry Specification Group (ISG) within ETSI that is working on MEC standardization. This initiative aims at developing MEC related specifications mainly the mobile network domain. Doing so, it works to unite the telecom and IT-cloud efforts to enable new applications at the RAN level. On the other hand, the OpenFog consortium, founded by high tech companies and academic institutions (Cisco Systems, Intel, Microsoft, Princeton University, Dell, ARM Holdings), aims at creating a reference architecture to apply fog in the IoT domain.

5. IoT security and privacy: an engineering perspective

The security and privacy issues hinder the IoT realization. Although, some of the IoT security breaches/vulnerabilities are common with the current Internet network [172], IoT presents new security concerns that make it the “Internet of Vulnerabilities” [173]. Some analysts argue that the security concerns in IoT outweigh its benefits. DDY intruder, DoS/DDoS, physical attacks, privacy attacks, eavesdropping, data mining, and traffic analysis are primary IoT attacks [174]. Additionally, new types of attacks related to the constrained things characteristics (low power, low processing, etc.) are IoT specific [175]. Such constraints expose devices to new type of attacks (running out of power, running out of memory, etc.) [176]. Thus, there is a need to propose security solutions that limit the effects of these attacks [177].

In this section, we will review the work done in the IoT security domain. Four main security aspects are considered: identity management, authentication, access control, and trustworthiness and privacy. At the end of this section, we will show how SDN/NFV (SDNv2) can be employed to overcome the security challenges in IoT and how it serves in the development of a security embedded architectural solution for IoT.

5.1. Identity management

Practically, what we refer to as being thing identities in IoT, are precisely things identifiers; the identities are more subject-related characteristics (in analogy with the human being case, the identity is the name, last name, birthday, etc.). Usually, in the online systems, we employ identifiers that are a set of uniquely identifiable strings [178,179].

As mentioned before, heterogeneity is one of the IoT challenges. One aspect of this heterogeneity is the presence of different identity schemes in the IoT domain. Device identification is straightforward with RFID; it is performed via the Electronic Product Code (EPC) scheme, which can distinctively recognize “things” from their tags [180]. In ZigBee, devices are identified by their network address, a 16-bit local unique identifier within one ZigBee network. ZigBee also assigns identities to networks via the Personal Area Network Identifier (PAN ID) and the Extended PAN ID (EPID), which are used to refine the identification process [181]. In Bluetooth, devices possess unique UUID identities that are hardcoded into them at manufacturing. The UUID relate to the Device Identification (DI) service record, and a device can have many DIs if it assumes many logical functions. In Wi-Fi (IEEE 802.11), device identification is based on a shared network identifier (SSID) at the access point level on one hand, and on the unique MAC address of the station on the other. As for UMTS, identification is based on the User Services Identity Module, which contains the permanent user’s identity (IMSI) and the temporary identifier (TMSI). Finally, in WSN, device identification shifts from being device ID centric to data content identification. The nodes may no longer be identified by their own IDs, but through the data they possess or require. This scheme is referred to as content-based names/addresses [182]. In the Internet, the devices are identified by their IP address (IPv4 or IPv6). In the telecommunication domain, the user equipment is identified by its embedded IMSI code, and the user has a phone number. A summary of the most known device identity schemes is presented in Table 5.

Therefore, having a unified identity scheme is critical to overcome the identity fragmentation in these vertical silos of networks. IPv6 is argued to be the most suitable solution for identification. IPv6 based protocols such GloWBAL, 6LoWPAN, IPsec, and MIPv6 are proposed as solutions to the power related and mobility challenges. Though IPv has been the Internet oxygen, it is not obvious that IPv6 will have the same role in the IoT domain. Already, IPv4
Table 5

<table>
<thead>
<tr>
<th>Domain of use</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4</td>
<td>Internet</td>
</tr>
<tr>
<td>IPv6</td>
<td>Internet</td>
</tr>
<tr>
<td>IMSI</td>
<td>Mobile Network</td>
</tr>
<tr>
<td>Mac Address</td>
<td>Internet</td>
</tr>
<tr>
<td>Bluetooth Address</td>
<td>Bluetooth Network</td>
</tr>
<tr>
<td>RFID-EPC ID</td>
<td>RFID tags</td>
</tr>
<tr>
<td>ZigBee-PAN ID</td>
<td>ZigBee Network</td>
</tr>
<tr>
<td>OneM2M identity</td>
<td>M2M Network</td>
</tr>
</tbody>
</table>

has been depleted and IPv6 adoption is still encumbered in the current Internet. Benefiting from the IP established protocols is key, but we might need to apply new architectural designs and management paradigms [183,184].

In [185], Zhi-Kai Zhang presents a new IoT naming scheme, proceeding from the IoT ITU definition that is based on the ability to connect anything at any time, from anywhere. There is a proposition of a property-aware name service (PNS). PNS mixes the “what”, “where”, and “when” aspects of the IoT definition in the object name conception. The object name provided consists of two parts: object name and location (Object_Name::Loc), both containing time stamp information. This scheme needs name resolution and location resolution servers (NRS and LRS), which resembles DNS. The N and L parts which provide time-validity checking are compared to the DNSSEC protocol for name based authentication mechanism. The overhead and delay added by the DNSSEC certificates queries make it unsuitable for real-time object name and location resolution. Besides the proposed scheme presents flexibility and interoperability and authentication facilities. In [186], there is a focus on the relation between things and users (owners). Therefore, upon connecting to certain device, you have to be permitted by the devices’ owner. In this case, the identity of the thing is related to its owner’s identity.

Friese et al. in [187] introduce the Kantara initiative’s IDentities of Things (IDoT) discussion group. The discussion group’s mission is to identify and analyze the main things identity related issues and to report the existing platforms used or proposed in this context. The authors claim that the name-based scheme (DNS) is not suitable to the IoT case. Regarding the authentication, they stress the importance of context-based authentication and concerning the authorization, they introduce the user-managed access protocol (UMA) on top of the access control framework OAuth.

The work done in [188] presents an IoT architecture and includes the most important technologies used at each level. OpenIDM is proposed for identity management, OpenAM for authentication management, OpenIG for authorization and OpenDG for data accessibility. JSON, REST, OAuth2, LW2M, DTLS, JS are some of the proposed technologies referred to in this work.

5.2. Authentication

Authentication consists of exchanging identity based information (or credentials) between two parties to confirm the identity authenticity. This service is intended to prevent masquerade and identity spoofing. Cryptography based methods have been established to perform authentication (one way and mutual authentication). However, strong authentication schemes invoke complicated cryptographic operations being computationally expensive. Applying these methods in the IoT domain encompassing a huge set of constrained devices is critical [189]. Some work has considered the emergence of IoT gateways able to handle the computational operations instead of the devices.

In [190], Turakovic et al. introduce a different perspective of IoT authentication, where the user and the node authenticate themselves directly and not through a gateway. The scheme was presented in a wireless sensor network context, where most of the nodes are of low performance and only few of them are gateway nodes (GWNs) with higher memory and capacity. The GWNs store IDs and keys of all other WSN nodes, and shared keys with users equipped with some sort of a smart card. The authentication step is launched only after the user logs-in and it requests a connection to the node directly. The two parties then share a secret key for subsequent exchanges. Mutual authentication for all three parties is needed in order to safeguard the key exchange session. This scheme is a lightweight yet robust scheme; the authors prove that it provides mutual authentication and key agreement, and security of all passwords. However, the GWNs are required to store IDs and passwords of all WSN nodes and all users, which constitutes a scalability issue for the method, especially that the GWNs are relatively limited in performance and memory.

In [191], a permit code authentication method is proposed. This method is lightweight and can be applied to constrained devices. In [192], a novel continuous authentication scheme is proposed. This scheme uses a public key scheme due to its efficiency in terms of scalability and memory use despite its computational overhead. The public keys are used to generate symmetric keys used as authentication token. The main concept that the proposed scheme introduced is the time factor. So, the generated key is a function of time and the two invoked parties can communicate over a certain period without having to pass into the authentication phase at each time they want to send/receive messages in a short time, which will reduce the overhead in terms of processing, delay and bandwidth consumption.

In [193], there is a proposition of an authentication scheme which relies on an asymmetric authentication method. The ECC algorithm is chosen to generate the private/public keys. In this method, things have their keys generated at the certificate authority (CA) via a secure channel, which is impossible in the ubiquitous IoT network access. The node has to know the public key of each node that it wants to communicate with, and a combination of the other node public key, its private key, and a random nonce are exchanged to do mutual authentication.

Kalra et al. introduced a new feature to cloud server/devices mutual authentication that relies on HTTP cookies and ECC. The protocol is divided into 3 steps. The first step is the registration, where devices subscribe to the server in the cloud by sending their unique identifier. The server would have chosen an elliptic curve, a point G on that curve and a private key. When it receives the identifier, it computes a cookie that is a hash of the unique identifier and the server’s private key, encrypts it using ECC and sends it to the device. Whenever a device wants to connect to the server, it sends a hash of the cookie, used by the server to authenticate the device. The server then sends a security parameter to authenticate itself to the device in order to establish the connection. Consequently, they decide on a secret key to be used to encrypt subsequent message exchanges. This is a new technique that relies on cookies and that is independent of device type. However, all devices need to support TCP/IP protocol and HTTP. Furthermore, the secret key is simply XOR-ed with the messages to encrypt them, which is a weak and breakable encryption technique [194].

The RFID technology motivated Kevin Ashton, the British engineer at MIT lab, to launch the Internet of Things term. However, this technology does not provide any kind of authentication and presents many security vulnerabilities. Many authentication schemes have been proposed as shown in [195]. The ECC is chosen for being the most convenient one. This paper surveys the RFID based authentication schemes in IoT in the aim of identifying the best schemes for healthcare environments. The authors compared the performance and the security robustness of different authentication schemes in the literature. The comparison was
done using elliptic curves over F (2163) for key generation. Results showed that all the studied schemes are prone to many attacks. However, three proposals [196–198] were judged to meet the minimum requirements for healthcare IoT applications. The lack of security problem in the RFID domain is tackled in [199]. The EPC code which is widely used in IoT and which is embedded in low power and constrained devices, does not use any cryptographic method, and codes are transmitted in plain text which exposes the authentication process to counterfeit attacks. A lightweight password generation based on XOR is proposed giving some level of security. The RFID technology proved to be efficient in IoT concerning object tagging and identification; especially that it supports all types of objects. Although it offers an edge to IoT, it suffers from many drawbacks, most importantly the lack of security. Aggarwal et al. study this issue in [20], showing advantages and disadvantages of RFID, and propose an improved RFID scheme for IoT. Their method performs authentication at the tag level. The reader sends its ID XOR-ed with a 128-bit random number R and then shifted by the weight of R. The tag used the received value to recover R, apply transformation on it and XOR it with its ID. The resulting value is sent via the reader to the backbone server that is able to authenticate the tag by recovering it from the received value and comparing it to the stored ID. A system is as strong as its weakest link. Therefore, in order to ensure a secure IoT network with RFID, security should be enforced even on the tag reading level. Furthermore, the authors show that the scheme is resistant to many attacks such as replay and disclosure [200].

Shivraj et al. review different techniques used for IoT authentication. They also set forward their own authentication process that relies on the One Time Password (OTP) technique developed with Elliptic Curves Cryptography (ECC). In their design, a PKG unit holds the IDs of all devices and applications in the network. At this stage, nodes acquire their public keys from PKG and compute their respective private keys. When a connection is to be established between an application and a device, the latter sends the ID of the node they wish to connect to the PKG. The PKG automatically generates the corresponding private key out of which it computes a one-time key. This key is sent to both nodes, which validate the connection by comparing the key with each other. The scheme was shown to be more efficient than other existing methods when it comes to the size of the key and the security robustness. The KDC does not store Private and Public keys of devices, it only stores their IDs. Consequently, hacking the KDC does not incur compromising the keys of all devices in the network. However, the OTP adds computational overhead since the KDC is required to compute a new one-time key every time a new connection between devices and applications is to be created [201].

The authors in [202] combined IoT concept with the Federated Identity and Access Management (FIAM) technique to address device authentication. The method was inspired from the web. It is composed of four agents: the device or the thing that was implemented with Arduino, the authorization server implemented in WSO2 that allows the creation of users and OAuth applications, the authorization tool to enable timely access, and the MQQT unit that supports plugins for authorization services. In this scheme, the nodes and the MQQT unit verify each other via the OAuth platform. The proposed framework is a compilation of many existing standards. This system is built with specialized components, which makes it more robust and secure. However, integrating these components might be problematic, which is why the authors faced many concerns during implementation.

The work done in [203] approaches the two-steps authentication scheme used in today’s business transactions. Instead of using a verification code sent to the mobile phone, the authors propose the use of a smart card for generating keys on the devices directly. Having credentials and keys at the same place and issued by the same party might have security issues. The proposed scheme tries to separate data and encryption keys. This method provides high security. The need to have smart cards makes this method impractical.

The authors in [204] present a dynamic adaptive authentication scheme for IoT (DAoT). This scheme switches between key establishment (KE), message authentication code and the TLS handshake based on the energy level of the constrained device. This scheme allows energy saving. An evaluation of the effectiveness of the proposed scheme is done using the Crypto++, a TestCrypt benchmark tool. The energy consumption is measured using an energy cost model that gives an estimation of the energy cost of each cycle. The results show that ideal amount of energy cost savings by DAoT. And even if the state of devices changes, DAoT can adjust the cost gap in stabilized state by feedback control scheme. Dynamic adaptation of the authentication method to the energy level of the device.

The architecture proposed in [205] lacks the autonomy of things, which is a main concept in the Internet of Things. Bai et al. revisit the issue of the integration of IoT in Cloud computing, in the aim of providing data online that can be accessed anywhere at any time. The architecture is composed of three islands connected via MPLS tunnels. The first island consists of users; IoT enabled smart card (ISC) per user, and readers. The ISC assigns a unique identifier to each user and transmits data to readers periodically. The information collected by the readers is relayed to a smart gateway that filters data and sends it to the authentication island. The latter is in charge of verifying the identity of users and integrity of data. The authors use X.509 version 3 certificates based on ECC to achieve authentication at four levels: user’s authentication, mobile device authentication, smart card and cloud server authentication. When authentication is complete, data is sent to the cloud server and stored in the cloud. This technique works with different IoT applications and devices and overcomes the protocol/vendor specific limitations. Furthermore, the different levels of authentication provide a robust security design. The ISC card is attached to users, however, this scheme might not scale if identities were given to all things [205].

In [206], Sungchul et al. propose an authentication scheme for RESTful web services in the IoT. This approach considers that each IoT object is presented by a unique URI. The REST being stateless presents some issues at the authentication level. The proposed method utilizes the ID-based encryption.

A comparison of the most used authentication methods: password, token, smart card and biometric is done in [207]. The comparison shows that although the biometric based one is the most secure one to authenticate human beings, it lacks applicability in the IoT domain. The smart card based method is the second secure one to authenticate its owner based on different applications. Then, the token-based authentication scheme is more secure than the password based one that is considered as the least secure one. The authors stated that there is a need to investigate the impact and the challenges in adopting any of the existing authentication schemes in the IoT domain. Table 6 summarizes the advantages and disadvantages of these methods.

5.3. Access control

Access control is a very critical part of the IoT security scheme. Guaranteeing authorized access to the collected data is an important task. Access control was mainly tackled in the web of things context [208]. Frameworks such as: OAuth [209], Shiro [210], and LDAP [211] have been proposed to manage the things roles description and access rights. Access management in IoT is a strongly related task to the identity management one. Having a unique identity, the thing can be granted access to the appropriate resources.
Thus, the gateway layer in the IoT architecture will play a key role in the identity and access management process [212]. However, the high scalability of the IoT network makes the discretionary access control (DAC) configured per user or device used in today’s web-based applications not suitable in the IoT case. This calls for alternative scalable solutions. Mandatory access control (MAC) which was used in operating system domains can be applied with the SDN integration in the IoT domain.

5.4. Privacy and trustworthiness

Privacy, a term related to Personal Identifiable Information (PII), is the ability to decide who can see our private data. Having our private assets connected to the Internet, privacy is a very important requirement in the IoT case. On the other hand, trustworthiness is a measure of how much a service or communicated data can be trusted. Having constrained devices prone to diverse attacks and third-party applications, a trustworthiness model is needed to protect our devices and data from a malicious exploitation.

In [213], a model, that relates trustworthiness to privacy, is proposed. This formal model can be used in an automated security framework to evaluate trustworthiness and guarantee the data privacy. In the RERUM consortium report [214], a model of the trustworthiness of IoT services is proposed. This model relies on a measure of reputation that helps in deciding if we can rely on certain service or not. The logs of the trustworthiness level can be used for notification in case of drastic changes (e.g., attack). In [215, 216], Nitti et al. consider the social Internet of things. In this case, the data provided by the users are considered as services that need to be trusted. Thus, their model consists of measuring the trustworthiness of the users by asking their friends. The distributed evaluation framework is shown to be able to isolate malicious peers. In [217], a reference architecture for improving security and privacy of IoT applications is proposed. The authors consider the IoT applications as being the entities that need to be trusted by the IoT devices and thus an evaluation of the users’ rating for these apps is presented.

5.5. SDN and NFV based security for IoT

SDN presents some security issues due to some of its characteristics such as the centralization one that makes it vulnerable to DoS attack per example [218]. However, its improvements in terms of providing an architectural based security solution over weight its disadvantages [219]. SDN provides a global view of the network thanks to the control centralization and thus the monitoring and consistency verification tasks become straightforward ones. Furthermore, the mitigation of some attacks (e.g. DoS, DDoS, etc.) become easier [220–223]. Thus, flow-based security schemes can be implemented dynamically at the network edge [224, 225].

In [226], the first network access control (NAC) using SDN through the use of multiple flow tables in the OpenFlow v1.3 protocol is proposed. It is shown that there is a reduction of 72% in terms of packets exchange compared to the captive portal approach and up to 80% reduction in terms of authentication delay.

Additionally, due to the virtualization integration, the isolation of flows become seamless. Therefore, the identity management on top of the heterogeneous identity islands needs a decision centralization guaranteeing the uniqueness of control. The authentication function was traditionally accompanied with central entities (PKI Servers, CAs, etc.) providing keys, authentication certificates, and related security services. The access control task is also accompanied with central authority managing the access permissions and roles. Above and beyond, SDN provides flexibility and agility in configuring and modifying security rules which makes this network evolvable and updatable. Dynamism of the IoT applications and services call for easily configurable security rules and policies [227].

Indeed, as summarized in Fig. 7, SDN and NFV bring novel benefits to the network security domain and especially in the IoT large scale network case. Having a global view of the network with periodically collected statistics, the SDN controller can detect abnormal behaviors and isolate the concerned flows or nodes thanks to the virtualization techniques. Moreover, SDN presents flexibility and dynamism in configuring the corresponding rules at the data plane level. Additionally, the network functions are softwareized and thus intelligence can be easily implemented at the controller level for intrusion detection. Therefore, having multiple applications that might modify the data plane rules, verification of rules’ consistency can be implemented to avoid network error.

5.6. Challenges

Applying SDN and virtualization in IoT brings its own security concerns: concerns about device bootstrapping, identity management, key management, and authorization. Device bootstrapping and key management should be standardized in the future to provide a common management interface to facilitate secure device configuration, thus enabling large-scale IoT deployment [228].

After revising the work done in the IoT security domain, we found that most of the work presents limitations. The proposed

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Authentication schemes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentication scheme</td>
<td>Advantages</td>
</tr>
<tr>
<td>Public/Private Keys</td>
<td>Scalability</td>
</tr>
<tr>
<td>Symmetric Key</td>
<td>Simplicity</td>
</tr>
<tr>
<td>Biometric</td>
<td>Simplicity</td>
</tr>
<tr>
<td>Identity Card</td>
<td>Robustness</td>
</tr>
<tr>
<td>Passwords</td>
<td>Simplicity</td>
</tr>
</tbody>
</table>

Fig. 7. SDN & NFV security benefits.
IoT architectures lack the integrated approach to security. Adding security solutions after the fact will be costly and has historically proven ineffective. Therefore, the high IoT scalability demands automated security solutions [229–231].

No matter how robust are added security measures, a secure infrastructure is a prerequisite. The underlying IP based infrastructure is by itself vulnerable, starting from the IP spoofing attack to more significant vulnerabilities [232]. In this context, applying SDN and NFV can alleviate several of the IoT security challenges. However, as SDN brings benefits in terms of security management, it poses new security challenges. The centralization of the intelligence, at the controller level, makes the SDN controller a single point of failure. Particularly, if the controller is hijacked, the attacker gains control of the network. Moreover, the controller is prone to DDoS attack where the switches can be maliciously programmed to flood the controller with OpenFlow packet-in messages. On the other hand, the switches can also be hijacked, and thus inconsistent rules can be added compromising the network availability. Additionally, SDN allows third party applications which makes the network prone to malicious application attacks. Thus, unauthenticated applications and northbound interfaces can employ the controller to compromise the network consistency and availability [233].

6. IoT Big Data: a management perspective

It is not about “things”; it is about data. Effectively, the IoT innovative value lays on a collected data foundation [234]. The IoT “Big Data” is about 3 V’s: Volume, Variety, and Value [235,236]. To an extent, we can say that there is no IoT without the sense of data. It is not about the size of collected/generated data; it is more about the diversity, heterogeneity, disparity of this data. Having the data shared between different entities poses security and privacy concerns [237,238].

Handling the IoT “Big Data” in a global IoT architecture is evident. In this context, the integration of data interoperability in a general IoT architecture is key [239]. Managing this data calls for advanced data technologies. Cloud computing related aspects are expected to play an essential role in this context [239]. Additionally, SDN is expected to improve IoT big data applications [240].

6.1. Cloud computing based IoT solutions

Different cloud-based IoT architectures have been proposed in the literature. A typical cloud based IoT architecture is proposed in [241]. This architecture consists of three layers: the sensor layer, the cloud central layer, and the application layer. A sensor bridge connects the sensors to the cloud. An IoT framework is proposed in [242]. This framework consists of three layers: device layer, central hub layer and cloud layer. Essentially, the device has two principal parts: the micro-controller (i.e. Raspberry Pi, Arduino) and the communication component which allows its connection to the network. The central hub layer presents a kind of gateway that is a middle point between devices and cloud layer. The cloud layer consists of three subcomponents: web server, web application and database. In [243], the authors consider IMS (IP Multimedia Subsystem) as being the solution to integrate IoT and cloud. Their proposed architecture consists of three layers: The IoT device layer, the IMS core network, and the cloud layer. Most of the needed services (Naming, communication, management, etc.) are supported by IMS. Mobile Cloud Computing is principally introduced to overcome the mobile devices incapability in terms of storage and computation [244,245]. Mobile cloud computing an integration of mobile and cloud computing domains. It is a platform where both storage and processing are leveraged to a third party the cloud outside the mobile phone [246].

Thus, the principles to build an IoT cloud system are summarized in seven points in [247]: enabling virtualization, enabling emulation and simulation of IoT units, enabling monitoring, dynamic provisioning, enabling softwarization, providing software-defined elasticity, and providing elasticity at the different levels. The proposed software-defined machine (SDM) consists of three hardware and software layers: vertical domain application and middleware, general purpose OS, and hardware layer. To meet these requirements, new cloud management means have to be introduced. SDN, which will be revisited in detail as an enabler technology for IoT in Section 7, has retrieved its precious management role in the datacenter and cloud domains [248].

Software defined units are proposed to be the base of an IoT cloud system. Integrating SDN in a cloud system provides elasticity, dynamicity, automated provisioning, policy-based configuration, fine-grained resource consumption, self-service model, and API encapsulation of IoT resources and capabilities. The main component is the IoT unit encapsulating functional (storage, computation, communication) and non-functional (security, configuration, quality) aspects. These fine-grained and modular units compose more complex components on demand [249]. In [250], SDN is shown to enable processing of the IoT data at the network level. Consequently, the number of packets sent over Internet to the cloud decrease.

Cloud networking and cloud inter-networking with interoperability across different providers and platforms are provided through an overlay layer of federation management. OpenDOVE is used as a cloud orchestrator in [251]. Thus, open source cloud networking tools such as OpenStack, OpenDaylight, and Open vSwitch are used to manage cloud systems [252].

6.2. Fog computing/MEC/Cloudlet based IoT solutions

The distributed set of mobile devices impose a geo-distributed set of data. Additionally, the IoT high scale makes the centralization of data a critical mission. Consequently, a distributed set of fog nodes is required.

The proposed fog model in [253] follows the fog tier architecture design (presented in Fig. 8) adding the notion of IaaS interface between cloud and fog and defining a PaaS programming model for the fog layer. The proposed model separates or differentiates logical and physical entities. The fog applications are not deployed at the fog nodes, despite the mobile fog application is an ensemble of processes running on the computing nodes in cloud, fog or end devices. The application design consists principally of APIs and events handlers supporting the main fog functions (receiving, sending messages/notifications).

The authors in [254] shed the light on some heterogeneous use cases that call for the fog computing application. The authors
choose different use cases where fog nodes characteristics (e.g. mobility) are different. The authors add a new dimension to the big data Vs, the geo-distribution one that calls for fog integration. In this context, a high-level fog software architecture is described. This platform consists of four components: the devices, the abstraction layer, the orchestration layer (sense, analyze, plan and execute) and a northbound APIs that connect the orchestration layer to the application layer.

The integration of fog and cloud is proposed in [255,256]. The proposed IoT fog-based architecture consists of: analytics layer, virtualization layer, reconfiguration layer, and hardware layer. The fog nodes can be reconfigured to meet the different applications requirements.

Therefore, SDN, presenting management facilities, can be employed to deploy fog/MEC/cloudlet nodes [257–260]. In [261], a software defined fog node based distributed blockchain cloud architecture for IoT is proposed. The presented architecture aims at implementing distributed security scheme at the network edge level. In [262], a software defined fog-based architecture is proposed. Similarly, virtualization helps in facilitating the deployment of the fog nodes [263–265].

6.3. Challenges

The cloud-related solutions handle the IoT Big Data management (storage, processing, and analysis). However, the IoT data need to be contextually analyzed to extract useful insights from this data. Consequently, data from different cloud/edge nodes need to be collected for this aim. While SDN can be employed for interconnecting the datacenters and the edge nodes, the interconnection of the data nodes needs special protocols and networking paradigm (e.g. Information Centric Network (ICN)).

Furthermore, the fact that data need to be stored at different locations, the decision where to put which data needs to implement intelligence at the application level. Moreover, the security of the communicated data and the management of the access control to this data is a challenge that needs to be considered in any cloud/edge-based solution.

7. IoT heterogeneity: a middleware perspective

7.1. IoT gateway

Owing to the massive heterogeneity in the IoT domain and the presence of vertically integrated domains and applications, the call for a gateway layer is crucial [266]. Thus, the IoT gateway has to perform multiple functions such as: protocols translation (NATing [267]), service chaining, security related functions (firewall, authentication, access control, etc.), data mining, QoS management, mobility and handover management, and routing and forwarding packets (Fig. 9) [268].

In [269], Datta et al. propose a OneM2M based IoT gateway. This gateway consists of a OneM2M middle node performing mainly three functions: data analysis, resources discovery, and device management. The authors propose a fog computing architecture based on the OneM2M standard. The vehicular networks use case is shown as a direct application, where the gateways are fog enabled and deployed on the Road Side Units (RSUs) providing the consumer centric services such as data analytics and semantics, and vehicles discovery and management. Besides, the IoT gateway plays an initial role for providing services and resources discovery. The Distributed Hash Table (DHT) and Distributed Geographic Table (DGT) algorithms employed in the P2P networks are used to discover neighbors and services in [270]. Applying cloud/fog computing paradigm to the gateway layer is key to handle data related services. BETaaS (Building the Environment for the Internet of Things as a service) is presented in [271]. The authors try to employ virtualization on top of the BETaaS gateways, so each VM can run certain applications. However, the need for high processing computations cannot be met by the edge nodes presenting power and computing resources limitations. Therefore, the migration to the cloud is necessary for certain complicated tasks. Benefits of edge mining reducing the traffic between edge and core are presented in [272]. Data trimming is one of the Cloud of Things (CoT) challenges. In [273,274], a smart gateway functional architecture is presented. Main tasks of this smart gateway are: collecting, preprocessing, filtering and reconstructing data into more valuable one, uploading only necessary data to the cloud, tracking IoT objects and sensors’ activities, tracking IoT power constrained nodes energy consumption, security and privacy of the data, and overall services monitoring and management [275].

Use of mobile phones as IoT gateways having the capability of transferring the data over wide area networks has been proposed in the literature [276,277]. In [278], Datta et al. have proposed a gateway having the role of translating the data requests/replies collected from the sensors and transmitted to the mobile applications. The access to data can be done in two ways: polling request to the gateway or registering to the gateway for notifications. The gateway is mounted on a Google Application Engine (GAE). The user is presented by its mobile application profile. Thus, the gateway functionalities, the scalability concerns, and the security issues have to be considered extensively. In [279], Datta et al. have proposed a mobile application to Connect and Control Things (CCT). This application uses the mobile phone to connect and control M2M devices through a gateway. This gateway translates the different technologies. The authors deploy the proposed application on a GAE where the messages are transported via HTTP. Future extension to support other messaging protocols is necessary. These messages contain SenML (Sensor Markup Language) based data. An SenML extension to support actuation messages is proposed. The SenML [280] is an ongoing standardization work that defines a standard format of the sensor measurements.

However, closeness and hardware dependence limit the gateway’s capability to support dynamic IoT network features. Virtualization migration from the ICT world to the network domain
will have an impact on the networking functions deployment [281]. Now, the current IP packets are processed by multiple middle boxes (e.g. load balancer, firewalls, etc.), if special forwarding paradigm has to take place. However, the middle boxes closed infrastructure incur complicated management and configuration tasks. SDN & NFV are coming to hide these complexities making the networking functions software based tasks, that can be deployed anywhere and on any hardware. Thus, the management and configuration tasks become easier and maintainable [282]. IoT gateways, meant to be deployed in big numbers, have to benefit from the SDN & NFV paradigm to be easily manageable. Intelligence and service chaining are other features acquired integrating SDN & NFV in the IoT gateway layer. The smart IoT gateways have to perform networking and data related functions. The integration of all functions in hyper-convergent smart boxes with SDN & NFV and cloud integration is proposed in [283]. Open vSwitch is proposed as being an intelligent edge in [284]. The proposition of an intelligent gateway is done in [285,286]. An extended MQTT queuing method is integrated into this gateway to support an enhanced QoS management mechanism.

With the introduction of these revolutionary technologies, service chaining becomes a straightforward task [287]. Dynamic network service chaining built on top of software-defined edges is tackled in [288]. These edges are deployed in a datacenter as software engines running on virtual machines. An emulation is done using Mininet as proof of concept; Pox (SDN controller) is used to configure the switches/routers edge nodes with the correspondent rules. In this setup, the authors use hybrid switches which support both SDN and legacy network functionalities.

In [289], an edge-computing platform for IoT gateways, called Paradrop, is presented. This platform is characterized by the dynamism, the management through OpenFlow, the supported APIs and security functions. In [290], the SDG-pro (software-defined gateways programming framework) for cloud IoT system is presented. In this framework, software defined gateways are provisioned and deployed dynamically on edge nodes by IoT controller units in the cloud. This approach allows the “everything as code” paradigm to deal with the IoT network dynamism and scalability.

The work done in [291] tackles the IoT gateway problems. The authors claim that today’s IoT solutions depend on closed application-layer gateways. The authors compare the today’s IoT application specific gateway to having a browser for each website, which is not an intelligible solution. Their proposed architecture consists of having a smartphone as an IoT gateway. The used communication technology is BLE where the smartphone is the master node and the peripheral things are the slaves’ nodes. Each slave node sends beacons periodically to notice its presence to the nearer master and the master has the role to establish the connection between slaves. The smartphone can forward the IPv6 packets from the peripheral nodes, if supported. If not, it has to act as a proxy to translate the different packets to IPv6 format. Several questions can be posed concerning the security, privacy, trust, user incentive, and reliability.

In [292], different approaches used for conceptualizing an IoT middleware have been presented. These approaches are compared in terms of the challenges that can overcome; these challenges are mainly: interoperability, trust, scalability, mobility, heterogeneity abstraction, spontaneous events, random topology, multiplicity, unknown data-point availability, security/privacy, actuation conflicts, bootstrapping, extensibility, modularity, and real-world integration. The analysis shows that there is no approach that can tackle and overcome all the challenges. Additionally, some challenges: trust, actuation conflicts, and bootstrapping, are not solved yet by any middleware approach.

The work done in [100] introduces the device-centric approach comparing it to the data-centric approach that relies on collecting and providing data without caring about devices ‘identity. IoT-A, SENSEI, FI-WARE, BETaaS, IoT6, etc.: all these projects aim to provide IoT cloud based architectures where the main scope is the data (data-centric). However, in [100], the purpose is to provide or deploy sensing and actuating cloud services. So, the user can provision services despite of asking for data (service-centric).

Table 7 shows the different approaches used in conceiving an IoT middle ware. These approaches can be categorized in four types: device centric (where the focus is on the device itself, so it has its own identity), user centric (where the device identity is related to the owner identity), data centric (where the data has to be identified), and service centric (where everything is served as a service (XaaS)). Each of these approaches has its advantages and disadvantages.

7.2. Challenges

Applying SDN and NFV to design an IoT gateway presents many advantages: programmability, management flexibility, configurability, etc. However, being controlled by an SDN controller, the gateway needs to communicate with the controller to populate its forwarding table which makes the controller a single point of failure. In case of connection failure, the gateway cannot operate in standalone mode for an extended period of time while maintaining correctness of the rules. This hints at the need for hybrid gateways that can operate in two modes (SDN and/or non SDN).

Furthermore, employing cloud/edge computing techniques for data management at the gateway level pose new challenges. The data distributiveness calls for new data-based networking paradigm. Taking the decision of which data need to be processed and analyzed at the gateway level and which need to be transported to the cloud is another challenge that calls for data classification and tagging at the device level.

8. IoT scalability: an architectural perspective

Having billions of things connected to the Internet in the future network, the network architecture needs to be rethought. Many IoT architectures have been proposed in the literature (Table 8). This situation is similar to have multiple remote controls (for managing different types of devices the DVD, TV, AC, etc.) all functioning the same way, but no one can replace the other [303]. For enabling the IoT wide deployment, we need a common agreed upon architecture as the case of the TCP/IP Internet architecture. The architectural diversity and heterogeneity and the absence of interoperability between these different architectures devalorized their utility [304]. SDN is intended to overcome this heterogeneity providing a common control layer on top of these different IoT architectural silos. In the following, we will review the most known IoT architectures and the most recent work applying SDN & NFV in a generalized IoT architecture. The layered representation of these different architectures is summarized in Fig. 10.

8.1. iCore

The iCore project defines three main levels in its framework: the virtual object level (VO), the composite virtual object level (CVO), and the service logic level. These levels aim at abstracting the heterogeneity at the physical object layer and provide cognitive services to ensure reliability [305].

As part of the iCore project, [306] presents a distributed framework for IoT. This framework consists of four modules embedded in IoT daemon: the virtual object layer (VOL), the composite virtual object layer (CVOL), the service layer (SL) and the security management (SM) module. This framework tends to provide interoperability between different IoT application domains. Mainly, each
Table 7
Middleware IoT approaches.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device-centric</td>
<td>The devices are identified and connected to the network.</td>
<td>Connectivity and security management</td>
<td>Scalability and heterogeneity</td>
</tr>
<tr>
<td>User-centric</td>
<td>The focus is on the device-user relationship; the device’s identity is based on the owner's identity.</td>
<td>Scalability and management facility</td>
<td>Lack of M2M type of communications</td>
</tr>
<tr>
<td>Data-centric</td>
<td>The focus is on data; the data will be identified, labeled, and classified without caring about user or device identity.</td>
<td>Big data handling and Interoperability</td>
<td>Added complexity (e.g. data classification)</td>
</tr>
<tr>
<td>Service-centric</td>
<td>The focus is on the services; services are provided to registered users upon request.</td>
<td>Scalability and interoperability</td>
<td>Reliability</td>
</tr>
</tbody>
</table>

Table 8
IoT Architectures Initiatives.

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Description</th>
<th>Partnership</th>
</tr>
</thead>
<tbody>
<tr>
<td>IoT-A [293]</td>
<td>A proposed IoT Architecture Reference Model (ARM) under the F7 European project.</td>
<td>Alcatel Lucent (Belgium, France), CEA (France), CTR (Italy), CSE (Greece), FhG IML (Germany), Hitachi (UK), IBM (Switzerland), NEC (UK), NXP (Germany, Belgium), SAP (Germany), Siemens (Germany), Sapienza University of Rome (Italy), University of St. Gallen (Switzerland), University of Surrey (UK), University of Würzburg (Germany), VDI/VIDE-IT (Germany), VTT (Finland).</td>
</tr>
<tr>
<td>IoT6 [294]</td>
<td>A 3-year F7 European research project for researching the IPv6 potential for IoT.</td>
<td>Mandat International (Switzerland), Ericsson (Serbia), RunMyProcess (France), University College of London (UK), University of Murcia (Spain), Vienna University of Technology (Austria), University for Applied Sciences Western Switzerland (Switzerland), University of Luxembourg (Luxembourg), KAIST (S. Korea).</td>
</tr>
<tr>
<td>iCore [295]</td>
<td>An IoT project aiming at abstracting the heterogeneity and representing the different user/stakeholders view. Cognitive context awareness, reliability, and energy efficiency are main goals of the conceived solution.</td>
<td>12 industrial partners from which 8 are leading ICT manufacturers (Alcatel, Bell labs, Atos, Fiat, Siemens, Software AG, Telecom Italia, Thales), 5 SMEs (Zigpos, Ambient, Arago, Innotec, M3S, Trilogis), 4 universities (Delft University of Technology, University of Surrey, University of Piraeus, KAIST), and 5 research centers (Create-Net, JRC, TNO, VTT, WuXi SensingNet).</td>
</tr>
<tr>
<td>OneM2M [296]</td>
<td>A service layer abstraction to overcome the vertical heterogeneity while ensuring compatibility with older M2M architectures.</td>
<td>8 of the world leading ICT's companies (Arib, ETSI, Atis, CCSA, TIA, TSDSI, TTA, TTC), 6 global fora and SDOs (Broadband Forum, CEN, CENELEC, Global Platform, Next Generation M2M Consortium, OMA) and over 200 companies in all industry sectors.</td>
</tr>
<tr>
<td>IoTDM [297]</td>
<td>IoT data Broker for oneM2M based infrastructure</td>
<td>Committed by Cisco, ETRI, Echelon, Technicolor (it is part of the Lithium ODL version).</td>
</tr>
<tr>
<td>FIWARE [298]</td>
<td>Provides a set of APIs to develop IoT applications ubiquitous, seCRet in the Internet-of-things with Location and context-awareness project to enable development of secure and assistant life applications.</td>
<td>Independent Open Community</td>
</tr>
<tr>
<td>BUTLER [299]</td>
<td>Provides a set of APIs to develop IoT applications ubiquitous, seCRet in the Internet-of-things with Location and context-awareness project to enable development of secure and assistant life applications.</td>
<td>INNO, Ericsson (Spain), Telecom Italia, GEMALTO, CEA, CWC, FBConsulting, ISMB, 1 Home Lab, ST, University of Luxembourg, K.U. Leuven, TST, Jacobs University, ZIGPOS, Maya Technologies, Banco Santander, Santander City Council, Tecnalia</td>
</tr>
<tr>
<td>COMPOSE [300]</td>
<td>Collaborative Open Market to Place Objects at your Service</td>
<td>5 Industrial partners (IBM Israel), INNOVA (Italy), U-HOPPER (Italy), CELLNEX (Spain), EVERYTHING (UK), 4 research Institutes (BDIGITAL (Spain), BSC-CNS (Spain), Fokus (Germany), Create-Net (Italy)), 2 universities (Open University (UK), University of Passau (Germany)), and one standardization body (W3C).</td>
</tr>
<tr>
<td>IEEE Project P2413 [301]</td>
<td>No new architecture but a high-level description to enable cross-domain applications and compatibility between different architectures.</td>
<td>Broadband Tower, Cisco Systems, Emerson, EPR, Finger Food Studios, Hitach, Honeywell International, Huawei Technologies, Infocomm Development Authority (IDA), Intel, Kaspersky Lab, Korea Eelectronics Technology Institute (KETI), NIST, Qualcomm Inc., Renesas, Rockwell Automation, Schneider Electric, Senslytics, Siemens AG, SIGFOX, STMicroelectronics, Toshiba Corporation, Wipro, Yokogawa Electric Corporation, ZTE.</td>
</tr>
<tr>
<td>TRECIMO [302]</td>
<td>Testbeds for Reliable Smart City Machine to Machine Communications (TRECIMO) is a project under the European Union’s FP7, Future Internet Research and Experimentation initiative.</td>
<td>EUR, TUB, Fraunhofer, CSIR, UCT, ABS, ESKOM, IZCAT</td>
</tr>
</tbody>
</table>
object must run this daemon and some layers can be omitted due to power and processing limitations.

The fact that an integration of the proposed layers has to be performed in the IoT devices, this architecture presents scalability and interoperability limitations. IoT solutions are already there, so imposing change in the present devices is an impractical solution.

### 8.2. 3GPP MTC architecture

Taleb et al. in [307] present the 3GPP MTC architecture. 3GPP is the 3rd Generation Partnership Project alliance grouping the Alliance for Telecommunications Industry Solutions (ATIS), the China Communications Standards Association (CCSA), the Open Mobile Alliance (OMA), IEEE and the European Telecommunications Standards Institute (ETSI), the Association of Radio Industries and Businesses (ARIB), the Telecommunications Standards Development Society (TSDSI), the Telecommunications Technology Association (TTA), and the Telecommunication Technology Committee (TTC) as organizational partners. This Machine Type Communication (MTC) initiative aims at introducing the M2M communication into the mobile network supporting initially the Human-to-Human (H2H) communication. The presented architecture consists mainly of three domains: the device domain, the network domain, and the user application domain. The device domain is where heterogeneity resides; this heterogeneity is in terms of supported communication protocols, device capabilities (power, processing, and storage), and supported security measures. The network domain is mainly the mobile core network (e.g. EPS in the LTE case) [308].

Kunz et al. in [309] present the main requirements, use cases, and key issues over the successive 3GPP MTC releases (10, 11, and 12). New features are intended to be added with new releases. Essentially, in the 5G era, the MTC surely has to be supported and new capabilities will be added as this new network provides revolutionary features [310].

### 8.3. OneM2M

OneM2M is an M2M based architecture aiming to provide an IoT middleware [311]. OneM2M intends to combat the fragmentation by implementing a horizontally deployed middleware service layer above the different vertical M2M silos networks and applications. Swetina et al. in [312] introduce the OneM2M standard. Essentially, the OneM2M architecture consists of three layers: Network Service Layer (NSL), Common Service Layer (CSL), and the Application Service Layer (ASL). These layers are presented by three types of entities: Network Service Entity (NSE), Common Service Entity (CSE), and the Application Service Entity (ASE). Five node types are included in the OneM2M functional architecture: Infrastructure Node (IN), Middle Node (MN), Application Service Node (ASN), Application Dedicated Node (ADN), and Non-OneM2M Node (NoN). These nodes are separated into two categories: CSE enabled and Non CSE enabled. These nodes essentially reside in two domains; The field domain contains the IN which presents the provider services and in the field domain reside the MN which is typically a gateway, the ASN which is oneM2M device, the ADN which is a constrained oneM2M device not presenting service providing capability, and the non OneM2M device which is normally a network device providing the underlying network services (location service, management service, and triggering service) [313].

In a myriad of M2M architectures and solutions, there is a need to retrieve a common middleware layer to combat the fragmentation and provide interoperability between these different silos. The global standardization initiative is launched in July 2012. Then,
the first version is released in December 2014 with ten specifications published online. These specifications tackle the main services provided by the CSL such as: registration, security, service charging and accounting, subscription and notification, discovery, group management, location, network service exposure and service triggering, application and service layer management, communication management, data management and repository, and device management [314].

Husain et al. in [315] describe how the OneM2M architecture is meant to use the underlying networks services (and more precisely when the underlying network is a 3GPP mobile one). Three services mainly can be provided by underlying networks through the network service entities (NSE): triggering, discovery, and management. One of the standardized infrastructures is the 3GPP one. Mainly an AE in the field domain has to have an IP connection with the AE in the infrastructure domain to establish a connection and performs one of the four operations: Create, Retrieve, Update, Notify, and Delete. This IP connectivity can be served by the 3GPP MTC network. The ASN in the field domain is similar to a user equipment in the user plane and the infrastructure node is similar to an SCS in the control plane. The MTC architecture presents some services that can overlap with those provided by OneM2M. Therefore, it is necessary to do the mapping between these services and one of them is the identification one. The used identifiers in the MTC case are the external ID (M2M-Ext-ID or MSISDN) and the Trigger-Recipient-ID of the target CSE.

The work done in [316] aims at integrating the lightweight management protocol (LWOMA) into the OneM2M architecture. The identity management and the object registration are two key parts in any IoT management scheme. Mostly, the authors approach IoT from a web-based perspective, using unique resource identifiers (URI) for identifying the things. Additionally, they extend the CoRE (constrained restful environment) capabilities to legacy devices and integrate the proposed scheme into the OneM2M standard architecture. However, URIs used in the web context present some limitations in the IoT case where we have big number of resources.

In [317], there is a proposition of OneM2M based smart city architecture consisting of things as ADN interacting with a gateway presenting an MN that aggregates data and bridge the non-smart things to the infrastructure node deployed in a cloud system. In a smart city domain, multiple gateways are deployed, and they have to be registered to the central smart city central cloud instance. The integration of the M3 data management framework [318] and crowdsourcing use case for smart city is discussed for providing smart services.

The work done in [319] tackles the issue of data interoperability. Most of the standardization efforts focus on one of the two aspects of interoperability: data and communication. The current version of OneM2M does not include the data semantic integration. The idea in this work was to integrate the data ontology concept in the OneM2M architecture.

8.4. IoT-A

The IoT-A architecture consists of seven longitudinal groups: device group, communication group, IoT service group, virtual entity group, IoT business process management group, and application group, and two transversal groups: security group, and management group as shown [320]. Thus, IoT-A provides an abstract architecture model and does not define in detail the main functionalities. This will conduct in different implementations posing interoperability issues.

8.5. IoT6

The IoT6 architecture focuses on three groups of the IoT-A architecture: the communication group, the business process management group, and the security group. In [321], the IoT6 architecture is presented. This architecture consists of six groups: the communication group, the resources and services group, the process automation group, the applications group, the management group, and the security group. This IPv6 based architecture employs the IPv6 inherent benefits (unique addressing, no need for NAT, etc.) and standards (6LowPAN, CoAP, GloWbal). It provides functionalities such as: mobility, multi-protocols interoperability among heterogeneous things, intelligence distribution, cloud computing and mobile phone network integration, ubiquitous access, and management capabilities. This architecture is not detached from previous IoT architectures (IoT-A, Fi-WARE, OneM2M, etc.) but it extends them. Focusing on the communication layer, it provides functionalities provided at higher layers in other architectures in a complex way. It complements the existing architectures supporting IPv6 to resolve the IoT identification challenge. This architecture consists of three domains: the IPv6 compliant and non-compliant things (the non-compliant things have to be connected to proxy or gateway) domain, the IPv6 local area network domain, and the IPv6 wide area network domain for connecting different LANs. The discovery service is provided through “directories”; these directories have multiple interfaces: JSON, DNS, etc. and have to be connected to a discovery core that applies the ontology principles to overcome the heterogeneity challenge.

The focus on IPv6 as IoT enabler is understandable in the identification context. However, the interoperability between different existent identification schemes adds complexity to the IoT ubiquitous network [322].

8.6. IoTDM

This module was integrated firstly in the Lithium OpenDaylight version. It consists of applying SDN to the OneM2M architecture. Having the data collected and analyzed by a central entity coping with the different access technology heterogeneity is key to enable the OneM2M deployment. This project consists of integrating a OneM2M core in the ODL controller. This core acts as an IoT data broker. The OneM2M core is connected to different devices using different protocols.

Thus, this project shows that OneM2M and SDN are two complementary concepts. In this context, employing SDN & NFV to deploy the OneM2M architecture will accelerate the IoT realization [323].

8.7. Software defined IoT architectures

“Is SDN the De-Constraining Constraint of the Future Internet?”[324]. While the current network technologies are considered revolutionary relative to what preceded, their endurance is limited by the rigidity of the current network infrastructure. In today’s network, the configuration is done through low-level policies configured manually (via CLI). SDN came to hide the management complexity and allow for innovative applications and network services to meet the IoT requirements [325, 326].

Omnes et al. in [327] discuss the benefits of employing SDN & NFV in a general IoT architecture. While SDN permits dynamic configuration of the data plane policies and rules, the NFV allows the virtualization of resources lowering CapEx and OpEx. The authors defined main requirements for a general IoT architecture such as QoS guarantee, common service layer, new access network mentality, and big data management.
A restful software defined IoT architecture is proposed in [328]. This architecture consists of several modules: northbound Application Programming Interfaces (APIs), southbound APIs, processor, and database. The southbound interfaces deal with different protocols: HTTP, COAP, etc. The control plane consists of the processor and the database where the nodes state and information are collected. The southbound interface, which interconnects the control and the application planes, is principally REST based.

Describing the usefulness of SDN to enable agility, flexibility, and dynamism to overcome the today's IT problems, Tadinada in [329] introduces the Freescale SDN products: VortiQa Open Network Director and VortiQa Open Network Switch. Two use cases of SDN OF switches are presented. The first use case is where Open vSwitch acts as an IoT gateway managed by a cloud based SDN controller. The second use case is where an Open vSwitch is mounted on eNodeB to offload data from the Evolved Packet Core (EPC) network, providing better user experience and decreasing OpEx and CapEx. The main functions assigned to the IoT gateway are: forwarding data between end devices, protecting devices from external attacks, providing QoS guarantee, authenticating and authorizing the end devices, transferring data in a secure way between gateways (IPsec/tunneling), and managing access control and queuing. Therefore, the eNodeB Open vSwitch aims to separate the voice from the data packets making the data packets not traversing the EPC network.

In [330], an SDN based architecture for home automation is proposed. Today, a big number of home devices are connected to the Internet. The management of these devices in a traditional way is unpractical and unviable in some cases. The authors propose the Majordom’home management platform. In the proposed architecture: CO is the connected object, coCO is the community of connected objects, VO the virtual object, and Avatar is the user representation to manage its VOs. The Internet Service Provider (ISP) plays the role of the Majordomo: a software that allows managing the user’s objects (i.e. the client home objects) through virtualization. The Majordom’home architecture consists of user manager, VO manager, coCO manager, network manager, and application manager.

Extending the work done in [330], Boussard et al. propose a generalization of the CO, VO, Avatar, coCO, and coVO definitions to any smart environment: A CO now is not just a home device connected to the Internet, it is an entity that can generate, receive or impact the data flow in the network, the VO is an abstract view of this entity, the coCO as before a community of connected objects, the coVO is a community of the virtual objects in other words is an abstraction of the coCO, and Avatar presents the manager of the CO through VO. The proposed SDN architecture consists of three horizontal layers and one vertical layer. The data layer consists of all NE and COs that can receive and generate data without performing any forwarding/routing functions, the control layer is composed of two sub-levels: level 1 consists of the network controller and the CO controller and level 2 consists of the coVO controller, and on top of these layers resides the application layer. The management layer consists of different managers (network manager, VO manager, application manager), all encompassed in the Operation Support System (OSS). The control, application, and management layers compose the “majordomo”. As proof of concept, they tested the proposed architecture with two Majordom’homes (Bob and Alice homes). Each one has an Open vSwitch to which the home appliances are connected. A coVO controller, residing at the ISP side, controls the Majordom’home gateways. Scalability, auto configuration, and security and privacy issues are to be tackled in future work [331].

In [332], a proposition of an IoT architecture that employs both SDN and distributed data service (DDS) is presented. While SDN is used to guarantee data agility, flexibility, and mobility handling, DDS is introduced for big data management. The publish/subscribe paradigm has shown its usefulness in the Wireless Sensor Network (WSN) case. This data-centric approach makes the data an addressable entity. In the IoT domain, this concept is needed because IoT applications and services rely mostly on the analysis of the collected data. This architecture consists of three domains: the M2M domain where a gateway connects the heterogeneous set of devices, the network domain that includes different access networks (3G, LAN, etc.), and the application domain that includes the IoT applications.

In [333], there is a proposition of a Software Defined Infrastructure (SDI) manager, which consists of two essential components: the cloud computing controller (OpenStack), and the network controller (FlowVisor)). The main roles of the cloud computing controller are the collection of the users’ descriptions and the management of the computing resources. On the other hand, the network controller has the role of managing the network resources, collecting network topology information, and interacting with the Open vSwitches to configure their forwarding tables. The FlowVisor layer is added to permit slicing of the network and the attachment of each slice to a certain controller.

The idea of Network Operating System (NOS) was depicted to hide the heterogeneity in the network domains. NOS allows the deployment of different applications over a set of different network devices. In [334], the authors propose an operating system for IoT extending the ONOS SDN controller to support SDN-WISE, a protocol that extends the SDN capabilities to WSN. SoftINTERNET a new initiative for a future software defined Internet. This architecture aims to provide both connectivity and management in a software defined way coping with the heterogeneity and complexity of the future Internet [335].

The trial to invoke SDN in the IoT domain is challenged by the delay imposed by the communication between switch and controller. A pre-emptive flow installation algorithm is proposed in [336]. In [337], there is a proposition of a software defined solution to overcome the heterogeneity challenge in the IoT networks. This solution consists of having an IoT controller which communicates with the things that have integrated IoT agents permitting them to request communication. These communication requests are collected by the IoT controller which builds a full view of the network and calculates the forwarding rules. These forwarding rules are communicated to the SDN controller which downloads them in the forwards (switches/routers). This solution builds an overlay network on top of the heterogeneous networks and allows the interworking between them. However, the proposed solution presents some limitations such as the integration of the IoT agents, the routing protocol, the forwarding rules formulation, the identity schemes heterogeneity, and the scalability.

Considering the scalability, management, and security IoT issues, the proposed architecture in [338] consists of three layers: the physical layer, the middleware/control layer, and the data service layer. The physical layer consists of different types of connected devices. The middleware/control layer consists of software defined blocks: Software Defined Security (SDLsec), Software Defined Storage (SDStore), Internet of Things Controller (IoT-C), and Software Defined Controller (SDN-C). When the data is received from the network gateway, a data collector process it; authentication is performed by the SDSec component, if the authentication check succeeds the data is tagged by a positive (P) flag (otherwise a flag N). Then, the data is passed to the IoT controller that has the role to compute the path to the destination, the forwarding rules are forwarded to the SDN-C which downloads them into the network switches.

Hu in [339] discusses the IIoT need for traffic engineering. Three phases of management are implemented in a centralized cloud based controller: topology computation, admission control, and allocation optimization. The centralization of management is shown
to have a good impact in terms of packet loss implementing an alternative route mechanism.

Employing the SDN concept, Lee et al. in [340] show that it is feasible to obtain interoperability between devices from different manufacturers. Auto configuration and recognition are integrated into the proposed solution. Using the Open vSwitch as a gateway, this architecture ensures a dynamic configuration and management of home networks. The home devices are identified using their unique MAC addresses. The configuration related information is kept in a database connected and managed by the SDN controller. An implementation of this architecture is done using Mininet and OpenDaylight with the home devices as hosts.

Two virtualization levels are defined in [341]: the network level and the end-user level. At the network level, there are two cases. The first case is where physical resources are in the same physical location; in this case, virtualization aims to partition the resources between different logical functions. The second case is where physical entities are at different locations; in this case, two virtualization functions are invoked: moving the logical function (migration) and having the physical resources at different places. Virtualization at the network and the end-user levels calls for specialized functions. The virtual sensor, the virtual cell management, and the software defined controlled wireless networks are presented as use cases that implement these virtualizations functionalities. Thus, SDN coupled with NFV can enable manageability in the IoT domain [342].

The work done in [343] presents a Web of Things SDN based architecture. The web technologies, facilitating the development mission, have some limitations in terms of security, things management (rebooting), and the data management. Therefore, putting SDN on top of the resource based Web architecture helps in hiding the security and management complexities. This architecture is composed of three layers: the access layer where the things are connected to WoT gateways, the control layer consisting of the resource databases and the control functions, and the application layer.

8.8. Challenges

Many proposals have been established for conceiving a widely adopted IoT infrastructure. However, the added functions complexities prevent their application. Even though the abstraction layers are promising, as the case of the OneM2M architecture, the techniques to deploy such layers need to be specified. In this context, SDN is a way to re-think the network functions deployment. Softwarization grants the dynamicity and the support of heterogeneity. The deployment of an SDN gateway is the solution to overcome the orthogonal diversity of IoT infrastructures. The reviewed architectures in this section show the benefits of applying SDN in facilitating the management network functions to cope with the high scalability challenge. Proposing a single central control can cope with the management issues. However, the centralization poses new challenges in terms of latency, availability, throughput, etc. Additionally, when we talk about a network with billions of connected things, the data management should be considered as well as the control layer design.

9. Limitations and future research directions

9.1. Limitations

Based on the review presented in this paper, we can list some of the limitations of the current IoT solutions:

**Lack of Interoperability:** different solutions have been proposed to overcome the different IoT challenges. However, most of them do not consider the existing IoT solutions and this makes the adoption of the new solutions a complicated task. Interoperability between the different IoT solutions (devices, architectures, protocols, etc.) helps in revealing the IoT value in enabling innovative applications. Therefore, in the network domain, adopting new solutions is not a straightforward task. Proposing a pure SDN based solutions per example is not realistic and thus the consideration of the hybrid case is key.

**Lack of Realizability:** scalability is one of the IoT main challenges. Most of the IoT challenges rise from the high scale of the IoT network which introduces new QoS and security issues. However, the realization of the high scale is not easy both from theoretical and practical perspectives.

**Lack of Compatibility:** having thousands of published papers in the IoT domain, few of them propose new schemes that can be integrated in the current network infrastructure. However, other proposals are meant to be standalone solutions that repose on new networking schemes.

**Lack of Security:** security is not an issue that can be treated independently. Security has to be designed and built in each layer of the IoT solutions (from the device layer to the application layer). IoT security is not only about securing the network and data it goes beyond that to attacks which can target the human health or life.

9.2. Recommendations

Based on the presented limitations, we believe that there are important directions that have to be considered in the future IoT research studies:

**Build on top of the existing solutions:** one of the most important challenges of any new proposed IoT solution is its interoperability with existing solutions. Thus, the future work has to consider the compliance of any new solution with the IoT standards and its interoperability with similar existing IoT platforms.

**Consider different challenges when building an IoT solution:** building an IoT solution for a specific challenge might result in a partial solution. Thus, it is essential for future work to define the main challenges that have to be considered in any IoT solution.

**Real implementation:** having plenty of position papers in the IoT domain, there is a need for real implementations that show the effectiveness of the proposed solutions. Additionally, real testbeds need to be implemented to test the correctness and effectiveness of the proposed solutions. Simulators that allow to model and test the different proposed architectures, protocols, and algorithms need to be developed.

**Consider data and network related aspects:** when IoT data related research focuses on data analysis the network related research focuses on how to connect things to the Internet. However, some applications present critical requirements in terms of network resources; thus, we need to rethink the data networking issue and where data analytics related functions have to be implemented.

**Standardization:** the interoperability between the different IoT solutions call for well-defined IoT standards. Therefore, the IoT standardization efforts at different levels (communication protocols, architectures, data management, e.g.) need to be correlated. Standards are key not only for compatibility and interoperability aims. Perhaps, the lack of standardization can invoke many security issues.

Actually, the IoT devices are not meant to communicate only with one device (i.e. gateway, switch, router). Instead, the IoT devices will communicate with many other IoT devices and thus standard D2D communication protocols are required. However, the standardization need to cover different IoT aspects and not only the communication one. The IoT devices will generate different types of data and thus interoperability at the data level is essential
to reduce data analysis complexity and enable innovative IoT applications. Besides, new regulations are needed to define the data ownership policies to protect data privacy and security [344].

**SDN, NFV, and cloud/edge computing integration:** as discussed in this paper, the application of SDN and NFV can alleviate many of the IoT challenges. SDN coupled with NFV provide flexibility and dynamism that help in overcoming the management complexity, aggravated by the high scalability of the IoT network. Additionally, SDN, enabling the programmability of the network functions, can cope with the IoT heterogeneity challenge. Furthermore, SDN and NFV can help in alleviating main IoT security concerns (e.g. DoS, DDoS, etc.). On the other hand, edge computing coupled with SDN can help in managing the IoT big data. However, the integration of these technologies need special consideration. In addition to the standardization efforts, developing real testbeds of the proposed solutions is key for their real deployment.

IoT involves many parties: IoT devices manufacturers, network services providers, data services providers, and applications developers. The IoT devices manufacturers need to monitor their devices for maintenance and management purposes. Additionally, some IoT devices manufacturers provide cloud-based services to store, process, and connect the devices’ collected data. In this context, SDN and cloud computing can hide the complexity of IoT devices and data management. Amazon, for example, has developed the Amazon Web Services (AWS) for IoT [345]. The AWS IoT services allow to manage the IoT devices through deployed applications in a cloud environment. In addition, it permits the management and the analysis of the collected data. On the other hand, the network services providers have to guarantee good QoS level and secure communication over their deployed networks. Thus, integrating SDN helps in managing both QoS and security in the highly scalable IoT network. Software based gateways help in overcoming the IoT network management complexity. Cisco, for example, has developed a software based gateway integrating Cisco IOS (i.e. Cisco IOS software with fog) to enable flexible IoT networks management and real-time IoT applications [346]. Besides, OpenDaylight has integrated the IoTDM module as a plugin ever since its lithium release. This module permits the connection of the devices, directly or through a gateway, to the SDN controller. In this case, the developed applications on top of the controller are responsible for managing the IoT devices and ensuring QoS and security. Furthermore, data service providers need to consider analyzing data at the network edge for this aim, for example, Microsoft has developed an IoT framework for implementing data analytics at the network edge [347]. Besides, application developers have to implement intelligence at the application level to get insights from the collected data. As an example, IBM has developed a cognitive system called IBM Watson for IoT data intelligence [348].

As a result, IoT data and network management requires the federation of all the involved parties’ efforts to have a global IoT architecture integrating the most recent data and network enabling technologies.

10. Conclusion

The network and telecommunication networks are in continuous evolution. Internet of Things is expected to take advantage of this evolution to be widely deployed. While some IoT applications are already there, its wide realization still encumbered by many challenges such as the high scalability and management complexity, the heterogeneity and interoperability support, the big data handling, and the security and privacy guarantee. These main challenges need new architectural and design based solutions to be handled. In this paper, we presented SDN & NFV, cloud and fog computing, and 5G as the main enablers of the IoT evolution. Many architectural solutions have been conceived in the IoT domain, but no one has gained a global acceptance and adoption. While we believe that SDN is a solution that combat the heterogeneity and can serve in unifying the vision of a global IoT architecture. While the work in this domain is still in its early stages, we presented the most recent work applying SDN & NFV in an SDN based IoT architecture.

**Acknowledgements**

Research funded by the AUB University Research Board, the Lebanese National Council for Scientific Research, and TELUS Corp., Canada.

**References**


78. M.H. Miraz, M. Ali, P.S. Excell, R. Picking, A review on Internet of Things (IoT), Internet of Everything (IoE) and Internet of Nano Things (IoNT), in: Internet Technologies and Applications (ITA), Wrexham, North Wales, UK, 2015, pp. 219–224.


Current for machines

Emerging

I.F. Oller Luo Trevizan Jin


Addison-Wesley

Jin1

5G, sheujp/lecture

and

software-defined

LoLo


H. Zhang, How to disconnect and secure the Internet of Things, Network Secur. 2016 (9) (September 2016) 18–22.


H. Zhang, How to disconnect and secure the Internet of Things, Network Secur. 2016 (9) (September 2016) 18–22.


T. Shon, J. Cho, K. Han, H. Choi, Toward advanced mobile cloud computing for the Internet of Things: current issues and future direction, Mobile Networks Appl. 19 (3) (June 2014) 404–413.


P. Hu, A system architecture for software-defined industrial Internet of Things, in: 2015 IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB), Montreal, Quebec, Canada, 2015, pp. 1–5.


Ola Salman received her M.E. degree in Computer and Communications Engineering from the Lebanese University in 2013. In September 2014, she joined the PhD accelerated track program in the Electrical and Computer Engineering (ECE) department at the American University of Beirut (AUB). Her research interests are in the area of Information Security and Networks, Software Defined Networks, Edge Computing, Artificial Intelligence, and Internet of Things. In 2017, she received the CNRS-I/AUB doctoral scholarship award from the Lebanese National Council for Scientific Research (CNRS) in recognition of her research work.

Imad H. Elhajj received his Bachelor of Engineering in Computer and Communications Engineering, with distinction, from the American University of Beirut in 1997 and the M.S. and Ph.D. degrees in Electrical Engineering from Michigan State University in 1999 and 2002, respectively. He is currently an Associate Professor with the Department of ECE at AUB. Imad received Best Research Paper Award at the Third International Conference on Cognitive and Behavioral Psychology (CBP), Best Paper award at the IEEE INFOCOM conference in June 2003, and at the International Conference on Information Society in the 21st Century in November 2000. Dr. Elhajj is recipient of the Teaching Excellence Award at the American University of Beirut, June 2011, the Kanal Salibi Academic Freedom Award, 2014, and the most Outstanding Graduate Student Award from the ECE Department at Michigan State University, April 2001.

Ali M. Chehab received his Bachelor degree in EE from AUB in 1987, the Master's degree in EE from Syracuse University in 1989, and the PhD degree in ECE from the University of North Carolina at Charlotte, in 2002. From 1989 to 1993, he was a lecturer in the ECE Department at AUB. He rejoined the ECE Department at AUB as an Assistant Professor in 2002, and became a Full Professor in 2014. He received the AUB Teaching Excellence Award in 2007. He teaches courses in Programming, Electronics, Digital Systems Design, Computer Organization, Cryptography, and Digital Systems Testing. His research interests include: Wireless Communications Security, Cloud Computing Security, Multimedia Security, Trust in Distributed Computing, Low Energy VLSI Design, and VLSI Testing. He has more than 180 publications. He is a senior member of IEEE and a senior member of ACM.

Ayman Kayssi studied electrical engineering and received the BE degree, with distinction, in 1987 from the American University of Beirut (AUB), and the MSE and PhD degrees from the University of Michigan, Ann Arbor, in 1989 and 1993, respectively. In 1993, he joined the Department of Electrical and Computer Engineering (ECE) at AUB, where he is currently a full professor. From 2004 to 2007, he served as chairman of the ECE Department at AUB. He teaches courses in electronics and in networking, and has received AUB’s Teaching Excellence Award in 2003. His research interests are in information security and networks, and in integrated circuit design and test. He has published more than 200 articles in the areas of security, networking, and VLSI. He is a senior member of IEEE, and a member of ACM, ISOC, and the Beirut OEA.