Radicand Economics

iMinds

dialogic
innovation • interaction

The impact of network virtualisation on the Dutch telecommunications ecosystem: An exploratory study

Commissioned by:
The Dutch Ministry of Economic Affairs

Project number:
2016.024

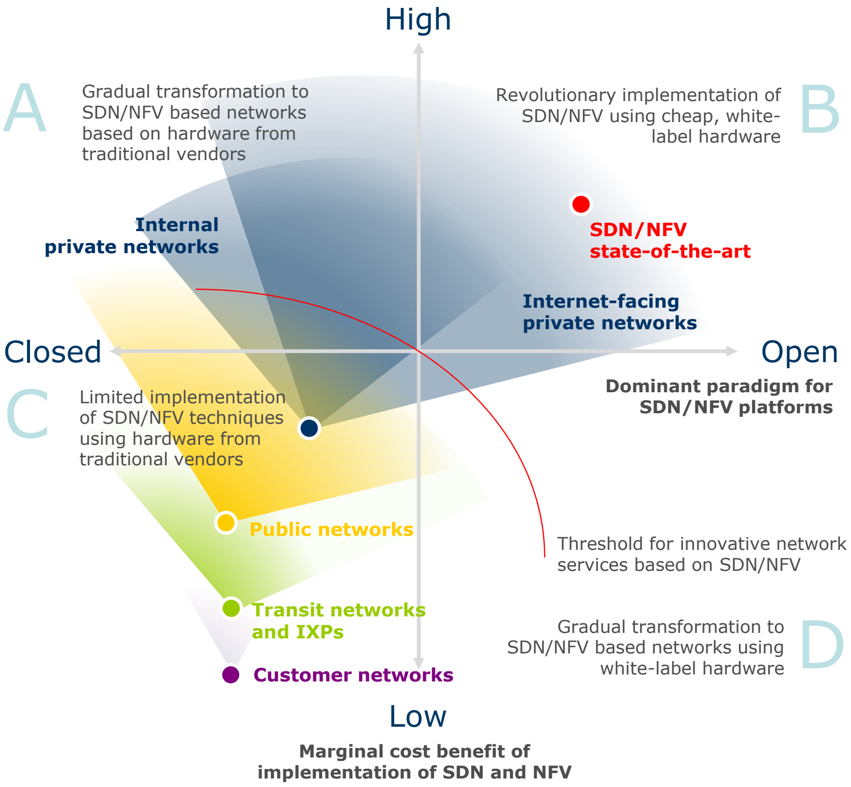
Publication number:
2016.024.1614 v2.24

Date:
Utrecht, 11 July 2016

Authors:
Tommy van der Vorst MSc
Bram Naudts
Paul de Bijl PhD
Prof. Sofie Verbrugge
Reg Brennenraedts MSc MBA

Management summary

Network virtualisation stimulates innovation by, from a technical point of view, enabling diverse network architectures to cohabit on a shared physical infrastructure. To make this possible, network elements need to be 'programmable' - software-defined networking (SDN) makes the control plane programmable, while network function virtualization (NFV) does the same for the data plane. By doing so, network architecture complexity becomes lower (e.g. easier network configuration via a centralized SDN controller), network-related expenditures are reduced (e.g. NFV promotes the usage of programmable general-purpose hardware), and it becomes less demanding to innovate (e.g. network operators become less dependent on standards development organisations and vendors to introduce new features).


High
A
Gradual transformation to SDN/NFV based networks based on hardware from traditional vendors
Revolutionary implementation of SDN/NFV using cheap, white-label hardware
B
SDN/NFV state-of-the-art
Internal private networks
Internet-facing private networks
Closed
Open
Dominant paradigm for SDN/NFV platforms
C
Limited implementation of SDN/NFV techniques using hardware from traditional vendors
Public networks
Threshold for innovative network services based on SDN/NFV
Transit networks and IXPs
Gradual transformation to SDN/NFV based networks using white-label hardware
D
Customer networks
Low
Marginal cost benefit of implementation of SDN and NFV

The implementation of SDN and NFV is driven by two forces in the market, visually depicted in the figure above. The first is the *dominant paradigm* for SDN and NFV, which is an attribute of the supply side. The second is the *adoption rate* for SDN and NFV technologies, which is an attribute of the demand side. Network operators are under pressure to come up with more innovative solutions, at least to control CAPEX and OPEX, and also to introduce new functionality and services to end-users, by using network resources in more flexible and efficient ways. In the longer term, without excluding the possibility that the market may remain stuck in scenario A, the market will therefore migrate to another scenario in the coming, say five, years. The degree and way in which this happens, is depicted in scenarios B, C and D, relative to baseline scenario A.

Policy recommendations

Competition

Analysis of efficiency drivers in relation to the different scenarios showed that network virtualisation is beneficial for static and dynamic efficiency, and will strengthen the positive externalities of ICT for the economy and society as a whole. Nevertheless, the impact on specific public interests, for instance related to cybersecurity, are uncertain, as new risks may come to the surface, while at the same time, networks and applications may become more resilient to threats.

By construction, scenario D (open paradigm, high adoption) sketches the most attractive perspective for welfare, both in the short and in the long run. The scenario analysis, however, is not able to establish the likelihood that a given scenario materializes. Similarly, the analysis is not suited to identify policy proposals that make scenario D more likely to come about.

Entry

From a technical perspective, the functionality offered by NFV/SDN is already possible by using currently existing technologies. While we do not foresee revolutionary business models based on the technical merits of SDN and NFV, we do expect evolutionary introduction and modification of business models based on the incremental improvements in organisational efficiency provided by SDN and NFV.

Network access

Direct (physical or logical) access to networking equipment will stay relevant for the years to come, due to the fact that not all functionality is available when an abstraction layer is used. Also, for debugging purposes, direct access remains relevant.

For now, it is important to first of all monitor market and technological developments. Physical access to networking equipment as well as access to lower layers (layer 2 specifically) appear to remain relevant for the next coming years as SDN and NFV mature. While this may not require access regulation to be changed just yet, we advise to investigate whether policy can be changed such that(in the future) access to (certain parts of) a network orchestration layer may be regulated.

Net neutrality

Without further clarification with regard to regulation of virtual networks, network operators will need to turn to the market authority in order to decide which services are provisioned ‘outside the internet’ and which remain on the over-the-top connection. If the decision remains at the discretion of the network operators, it might, at least in theory, lead to additional barriers to entry for service providers.

To guarantee the aim of net neutrality whilst at the same time having market parties and society reap the benefits of SDN and NFV, a (continued) dialogue between regulators and market parties is advisable.

Security & privacy

The implementation of SDN and NFV and centralisation of control may grow the attack surface of network infrastructure. Centralisation of control may also provide a way to easily audit the security and traffic policies in a network for compliance with rules and regulations (e.g. on privacy, net neutrality. The abstractions provided by SDN and NFV in theory make it possible to swap hardware from one vendor with hardware from another. At this point we do not see reason to adapt policy specifically to address security or privacy concerns following from the adoption of network virtualisation technologies.

Table of contents

1 Introduction ... 7

1.1 What is virtualisation? ... 7

1.2 Relevance for society ... 11

1.3 Research questions ... 13

1.4 Methodology ... 13

1.5 About the researchers ... 14

1.6 Reading guide ... 15

2 The technology push ... 17

2.1 The internet of today ... 17

2.2 The challenges for network operators ... 25

2.3 Network virtualisation: a way out? ... 29

2.4 Overview ... 35

3 Implementation scenarios ... 37

3.1 The different playgrounds for SDN and NFV ... 37

3.2 Opportunities for SDN and NFV ... 43

3.3 New propositions ... 45

3.4 Implementation drivers ... 50

3.5 Implementation scenarios ... 53

4 Economic impact ... 57

4.1 Economic concepts ... 57

4.2 Technological change as a primary force ... 60

4.3 Developments and economic impact ... 61

4.4 Overview ... 73

5 Conclusion ... 75

5.1 The road ahead for SDN and NFV ... 75

5.2 Answering the research questions ... 77

References ... 83

Appendix 1. List of interviewees ... 85

One-on-one sessions ... 85

Workshop session ... 85

1 Introduction

Since the introduction of digital telecommunication, networks have been growing and increasing in capacity exponentially, following the extraordinary developments in semiconductors and the larger computer industry. The growing capacity and also complexity of networks has lead the industry to create technologies to make deploying, managing and monitoring networks an easier task. Amidst the rapid developments, network vendors have to keep up with competition, and do not want to be bound by the relatively slow development cadence of hardware. Instead, the telecommunications industry have taken a page from the computer industry book, and embraced the practice of *virtualisation*.

1.1 What is virtualisation?

While it is obvious today that information technology (IT) hardware is general purpose and usable for different tasks, this has not always been the case. In 1971, when Intel introduced its first '4004'-microprocessor, their (dominant) competitors described it as a chip to operate traffic lights. Semiconductors were primarily considered to be an invention that could create more efficient and lower cost versions of existing, mechanical automation solutions. At that time, microprocessors were 'single purpose' and used for a specific task [16].

Since then, storage, compute and network resources have become exponentially faster and more compact while component costs have also decreased exponentially (Figure 1). This exponential growth is in the first place driven by the developments in the semiconductor industry (Moore's law).

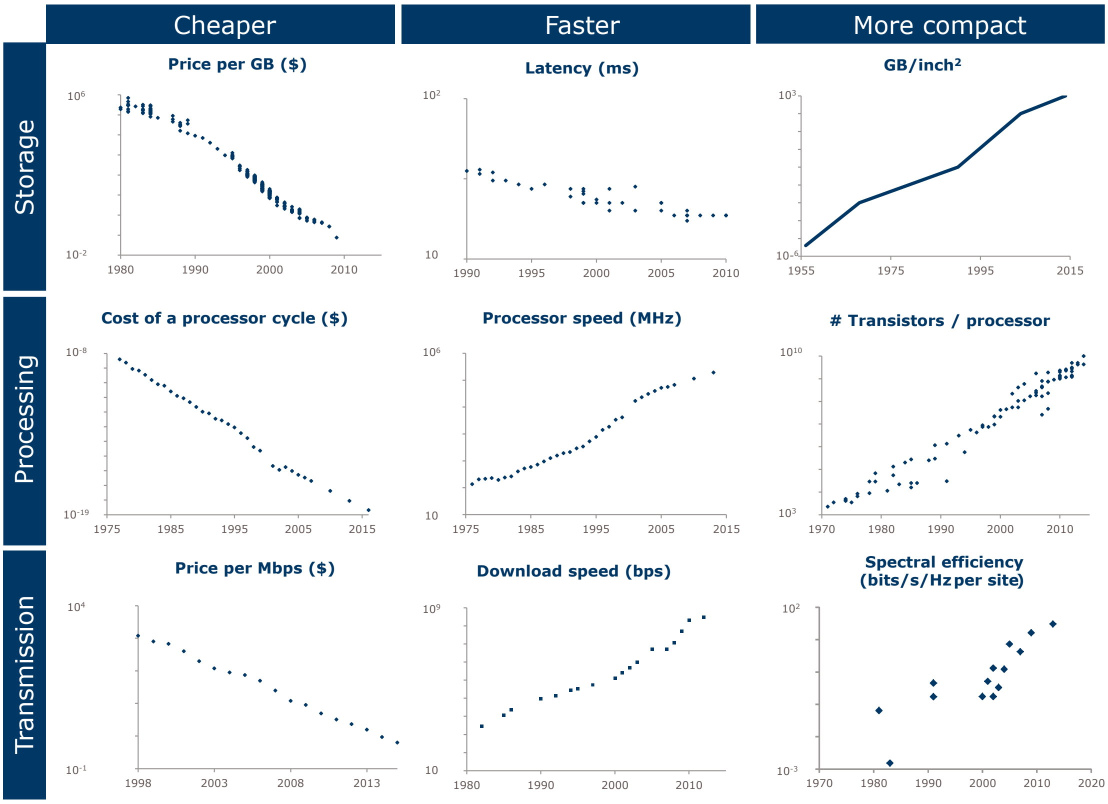


Figure 1. Exponential growth of the three core components of ICT: storage, processing and transmission of data [16]

Virtualisation refers to the creation of a virtual version of a particular hardware or software component. For the user, the virtualized version acts exactly as the original, non-virtualised underlying component.

In the context of networking, virtualisation refers to the practice of abstracting away low level networking hardware and software components into a single, virtual, software-defined entity. With virtualisation, networks are no longer defined by the sum of physically interconnected hardware boxes and cables, but are completely virtual entities, fully defined and managed in software, and running on flexible, replaceable hardware.

1.1.1 Three types of virtualisation

Virtualisation is an enabler of both process and product innovation. With virtualisation, the same (general-purpose) information technology building block can be used for multiple applications, leading, in many cases, to efficiency gains compared to single purpose solutions (*virtualisation for efficiency*). Virtualisation also permits flexible use of IT (*virtualisation for flexibility*). Finally, virtualisation allows to add intelligence to IT (*virtualisation for intelligence*). Figure 2 illustrates these forms of virtualisation schematically. The three forms of virtualisation are described in detail hereafter.

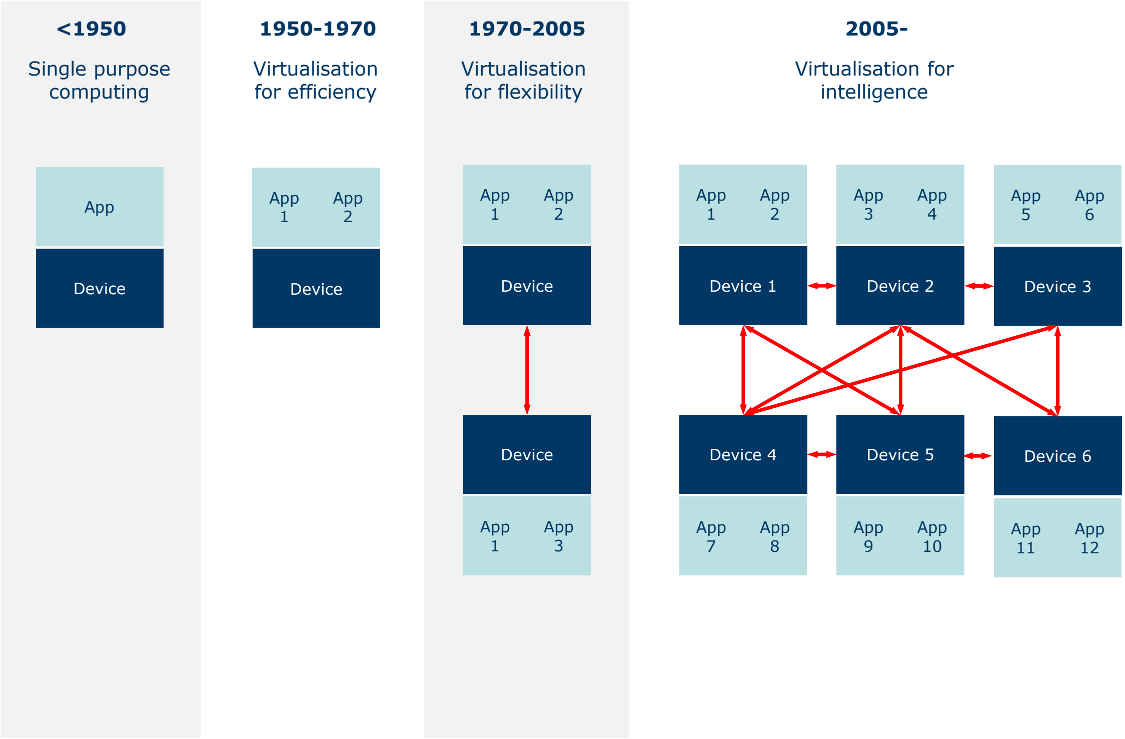


Figure 2. Four phases of virtualisation: evolution from no virtualisation ('single purpose') towards intelligent infrastructure.

Virtualisation for efficiency

The exponential growth of storage, computing and network capacity leads to a situation in which the capacity provided (supply) may, in many cases, exceed the capacity used by a particular application (demand). For example, a web server hosting a simple website will have ample free resources for most of the time it is turned on. In addition, usage patterns of applications often show a high level of fluctuation over time. For this reason, telecommunications networks are always overprovisioned.

The availability of free capacity which may be a magnitude larger than the used capacity introduces the possibility to consolidate several applications: instead of running each application on its own hardware, hardware can be shared between multiple applications. For example, one web server can be used to host tens or hundreds of websites.

To avoid the situation where one malfunctioning application would disrupt the normal behaviour of applications that run on the same physical hardware, virtualisation technology is used to create isolated virtual versions of the underlying hardware. To the user, the virtual version is presented as if they are the only user of the hardware while in reality, the physical hardware resources are shared.

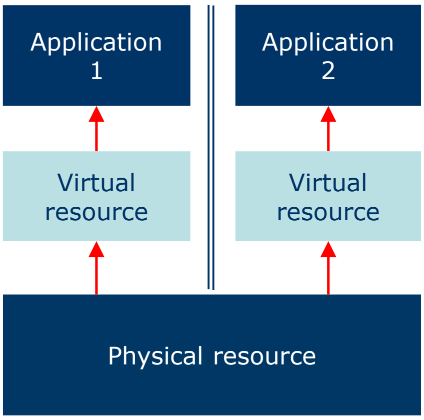


Figure 3 With virtualisation, applications use separate virtual resources.

Virtualisation for flexibility

The further proliferation of application areas for information technology drives the demand for flexibility of the underlying technology. By applying virtualisation one is able to meet that demand. Virtualisation allows to create a pool of physical resources that can be consumed on demand. For example, storage disks can be pooled to form a storage system which is flexibly allocated to the departments of a company, based on the department' demand.

In general, a virtualised system can adopt more easily to external changes to the system. For example, critical applications require high uptime (e.g. 99.99% or 99.999% uptime). A server crash, a power failure or a human error can cause a system to become unavailable. The benefit to using virtualisation is that it becomes simpler to reach (very) high availability (when used in combination with fault-tolerant hardware) as applications can move back and forth as needed between servers.

Virtualisation for intelligence

The last form of virtualisation enables the underlying infrastructure to decide autonomously on the services that need to be provided. This form of virtualisation makes it possible to orchestrate storage, computing and network resources in such a way that they are optimally allocated between applications. An example of virtualisation for intelligence is the way virtualisation is used in some cases to provide *high availability services*. When a hardware failure is detected on a system running a particular software component, the running software can, without interruption, be transferred to a different physical host. Transferring running services like this also requires that storage and network connections are moved along, in concert. This is schematically displayed in Figure 4.

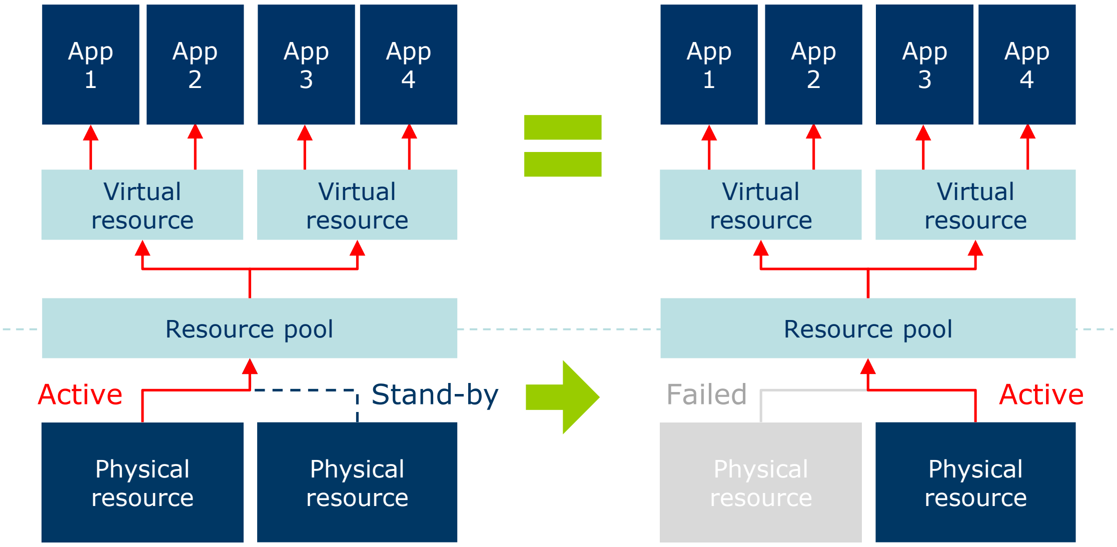


Figure 4Virtualisation for fail-over

1.1.2 Virtualisation of networks

A common misconception is that network virtualisation is a completely new concept. In fact, network virtualisation is as old as digital networks themselves. Virtualisation has been applied on virtually all levels of the networking stack. On the application layer, there are many applications that create over-the-top virtual networks, such as Skype, BitTorrent and Bitcoin. Further down into the transport layer, virtual private network (VPN) technologies and network address translation (NAT) are examples of virtualisation technologies that are in widespread use today. At the data link layer, VLAN tagging is a technology that is often used to create logically and functionally separate networks sharing the same switching hardware. Part of the whole raison d'être for these layers in the networking stack is that it allows one layer to be swapped out in favour of another, possibly virtualised, one.

Today, virtualisation technology for storage and computing resources is considered to be relatively mature. The abundance of mature technology has been a key driver for the market of cloud storage and computing. Several companies for example offer on-demand network access to a shared pool of configurable storage and computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Network virtualisation technology is however not yet as mature as virtualisation for storage and computing resources. Unlike in virtualised computing and storage, it is (usually) not possible for a customer to receive network resources on-demand. For example, a subscriber cannot simply ask its internet service provider to upgrade its up and downlink bandwidth from 10Mbps to 1Gbps and expect to have it changed instantly. Similarly a company cannot receive a network connection with a pre-defined quality-of-service (e.g. end-to-end delay of 50ms) on the spot. This has several reasons. First, the physical network resources may not be capable to provide the demanded bandwidth and it may not be straightforward to do so in a short term. For example, the physical carrier (e.g. copper) may need to be replaced by another one (e.g. fibre) or the distance the signal has to travel may be too long to provide sub 50ms end-to-end delay (e.g. satellite communication).

A second (and equally important) reason is that networks are complex to operate. As such, careful planning is required to provide a new service. Network virtualisation technology is considered as a promising enabler to tackle the second challenge. More and more heterogeneous devices become connected with each other under the Internet of Things paradigm. Examples include household appliances (television sets, thermostats, etc.), cars, et cetera. The combination of heterogeneous services and devices puts very diverse requirements on the underlying infrastructure. As such, the underlying infrastructure should be able to provide intelligently a number of storage, computing and network resources.

1.2 Relevance for society

The ongoing process of virtualisation in the computing industry has enabled the industry to move faster and deliver more flexible and efficient solutions, leading to significant growth, and also opportunities for the Dutch software sector. [16] It is expected that further development of network virtualisation technology will be able to deliver similar benefits. Network virtualisation also provides new challenges for policy, which needs to be adapted to fit with the new opportunities provided by network virtualisation.

1.2.1 Market perspective

The most important reasons for looking into the economic impact of network virtualisation are the following:

- **Network virtualisation technologies may enable new business models.** Three types can be distinguished:

    1. Market parties may start offering virtualisation hardware or software for management of software defined networks or virtualised network resources. Examples of this type are the offerings of companies like VMWare, Cisco and Juniper.

    2. Develop new network services that can be deployed atop of virtualised network hardware.[1]

    3. New business models of a lesser technical nature, e.g. verification of network configuration and policies based on network virtualisation technology.

- **Network virtualisation technologies may change the interrelations and power balance between market players in the telecommunications ecosystem.** The introduction of WhatsApp was a giant power shift from the operators to a single application service provider at higher network layers. Network virtualisation technologies may on one hand provide more control to higher level operators over lower level network elements, but on the other hand strengthen the position of the operators of the lower level networks in the first place, as they have full control over what kind of external control is allowed and what is not.

- **The interrelationship between the network equipment vendors and the network operators may change.** Virtualisation may enable operators to choose between a larger number of lesser specialised, general purpose hardware. Virtualisation may lead to more competition in the market for general purpose

[1] An example of such a service is Zerotier, which allows the creation of over-the-top Ethernet (layer 2) networks. Another example is the Fastly content delivery network, which is based on virtualisation technology.

solutions. It is however questionable whether this will also happen for more specialised solutions.

1.2.2 Regulatory perspective

There are significant differences between network virtualisation and the earlier virtualisation processes as they have occurred and are occurring in storage and processing. The primary difference is that networks are inherently bound to a particular location: they serve to bring traffic from one place to the other. This characteristic is defining for the telecommunications ecosystem and market competition. It has lead governments to create regulations specifically for telecommunications networks, of which no parallel exists elsewhere in the computing industry. The following regulations may be of relevance in the context of network virtualisation developments:

- **Net neutrality.** European net neutrality policy[2] prohibits internet service providers from providing internet service that favours traffic of one service over the other, except in very specific cases (e.g. for network management and congestion control).

  The aim of net neutrality policy is to guarantee that end users are able to equally access all content available in the market. Without net neutrality regulations, an ISP can enter into a contract with one content provider to favour its service over other services, posing an entry barrier for newcomers and competitors. In addition, an ISP may start to offer content services of its own and favour them over competition. Under net neutrality policy, content providers are able to compete freely without having to negotiate with the ISPs.

  Network virtualisation technologies may make it easier to flexibly configure network policies and create specific networks for specific applications on demand, which could have an effect on net neutrality regulations.

- **Unbundled access.** Access regulations require operators of fixed access infrastructure which are deemed to have 'significant market power' are required to allow other network operators access to (parts of) their infrastructure. In the Netherlands, the regulation applies to incumbent operator KPN, who owns the national fixed telephone network and is also the indirect owner of many fibre networks. Several operators, such as Tele2 and Online make use of the infrastructure of KPN (usually the 'last mile' between the customer and the first point of presence of KPN) to provide their service.

  Network virtualisation may provide new ways of providing access to infrastructure, which could create opportunities for setting access policy that can be executed more efficiently and fairly. Still, such a change would probably require replacement of existing network nodes.

- **Privacy and security.** The Dutch telecommunications law contains policy regarding the security and privacy on networks. For example, operators of public networks are not allowed to use so-called *deep packet inspection* to learn the contents of transmissions of its customers. Network operators (and more broadly, operators of ICT infrastructure) are also required to adhere to rules regarding the handling of

[2] EU Directive 2015/2120 was enacted in November 2015. [35] The Dutch net neutrality policy (defined in [36]) will be replaced by the European policy.

personal details and sensitive information, and are (since 2015) required to report any data leaks they may have experienced.

Network virtualisation may provide network operators with more powerful and centralized control over networks, which may require the regulations to be changed. Also, the centralized control of networks may turn out to be an interesting attack surface for hackers.

1.3 Research questions

Network virtualisation can have great influence on business models, market shares and common practices in the telecommunications industry. This has evoked questions at the Dutch Ministry of Economic Affairs (EZ): what is the influence of network virtualisation on the market and its ordering? Is the current regulatory framework appropriate for a future where many or all telecommunications networks will be virtual?

The Ministry has requested Dialogic to execute a short, exploratory study, to contribute to an orientation on impact and effect of virtualisation of telecommunications networks. In this study, we answer the following research questions:

1. Which new (business) models are enabled by network virtualisation for market parties as well as society? Which applications are currently foreseen, and what use cases and business cases are regarded by organisations who have been adopting virtualisation technologies as the most promising in the short term?

2. How will virtualisation change the competitive landscape of ISPs, network vendors and service providers, and how will it change their stance towards standardisation of, research & development (R&D) on, and deployment of recent network virtualisation technologies?

3. What are the *points of control* in virtualised networks, and which market parties will have access to these in the future?

4. Which types of regulated network access are needed to allow for effective competition and market entry of alternative service providers, who do not own an access network?

5. What influence can virtualisation have on net neutrality, and how should this be monitored from a regulatory point of view? In particular, what role can network virtualisation technologies play in improving *quality of service* (QoS) aspects?

6. How will recent network virtualisation technologies influence the security and privacy of network communication? Is there a need to change or update regulation regarding privacy and/or security requirements?

In this study, we consider a time horizon of five years (2016 – 2020). Also, the research questions are scoped to the Dutch telecommunications ecosystem.

1.4 Methodology

The primary goal of our study is to explain how the ecosystem for telecommunications in the Netherlands can evolve following the adoption of new network virtualisation technologies. In our experience, prospective technology studies are often highly biased towards static arguments, while in reality the market is moving much more dynamically. Disruptive innovations, such as the introduction of *over-the-top* services for telephony and messaging, can rapidly destroy existing business models of traditional network operators. Likewise, the role of network vendors can tilt in a very short period of time.

Clearly, predicting the future in such a highly dynamic market is challenging. In addition to technological uncertainty, there are also various competitive and strategic considerations that need to be taken into account. New opportunities for market entry will appear for new types of market parties, and the nature of strategic interaction between market players will change. Information asymmetry in the market on the demands and needs of end users, as well as the possibilities on the supply side, is another aspect that is of relevance.

In order to draw meaningful conclusions with respect to the research questions, we perform an *explorative scenario study*. In this study, we answer the research questions by starting from a clear and detailed view on the current telecommunications ecosystem in the Netherlands, especially regarding *ownership of networks*, *supplier power* and *buyer power*. After analysis of network virtualisation technologies, we sketch different scenarios that reflect the situation after five years of adopting new network virtualisation technologies. Figure 5 gives a schematic overview of the research methodology.

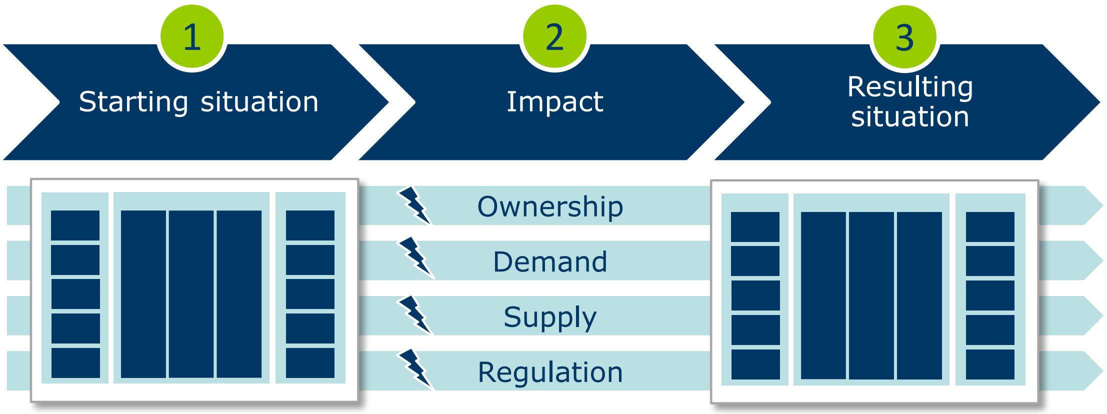


Figure 5 Schematic overview of the research methodology

The three phases of our research are the following:

1. Analyse the current position of the Dutch telecommunications ecosystem.
2. Analyse the direct impact of network virtualisation (first-order effects) on the Dutch ecosystem for telecommunications.
3. Sketching of different possible end positions of the Dutch telecommunication ecosystem in the future (scenarios).

In each of these steps, we consider four specific aspects: (1) ownership of networks, (2) demand for network services, (3) network equipment suppliers and (4) regulation. Note that in our study, we assume regulation is not changed during the time period analysed. Rather, we aim to provide insight in where the shoe will pinch in the future given current regulation, which may provide concrete pointers on how regulation could be improved.

1.5 About the researchers

Dialogic has been involved in public ICT and telecommunications policy for more than 18 years. Dialogic has performed numerous research-, consultancy-, and implementation projects for a wide range of customers, primarily in the (semi-)public sector. We have frequently performed prospective studies, both inside and outside the telecommunications domain.

iMinds is the hub for digital research and entrepreneurship in Flanders. iMinds conducts strategic as well as applied research on digital technologies. The *Internet Based Communication Networks and Services (IBCN)* group at the University of Gent, is part of iMinds, and consists of about 150 researchers working on various topics, including network virtualisation.

Radicand Economics provides companies, policy makers and regulatory bodies with consultancy on market ordering, competition policy, regulation and monitoring. Paul de Bijl, owner and founder of Radicand Economics, is specialised in topics at the crossroads of businesses and governments, specifically in the domain of telecommunications and ICT.

1.6 Reading guide

In chapter 2, we will first provide an overview of network virtualisation from a technological perspective. In chapter 3, we analyse scenarios for implementation of these technologies in the Dutch telecommunications ecosystem. In chapter 4, the different scenarios are translated to economic impact. Finally, in chapter 5, we discuss our findings and provide answers to the research questions.

2 The technology push

In this chapter, we analyse the recent push of technologies that has sparked renewed interest in network virtualisation from the industry as well as policy makers. This chapter provides the necessary background to grasp the technical and economic rationale behind network virtualisation. As we will see, specific technical implementation details may be of great importance in determining the effect the implementation of network virtualisation technology will have on society.

As the internet's protocol suite (TCP/IP) forms the basis for many of the (even private) networks in operation today, we start in paragraph 2.1 by introducing the ever-evolving architecture of the internet. In paragraph 2.2, we outline the challenges of internet service provider (ISP) and data centre (DC) network operators that are the consequence of the internet's dynamic nature. Finally, in paragraph 2.3, we discuss how network virtualisation technologies can help network operators in solving these challenges. In this paragraph we discuss the implications of network virtualisation and how it relates to software-defined networking (SDN) and network function virtualisation (NFV) from a purely technological point of view.

2.1 The internet of today

The internet is a large, essentially global, system of interconnected computer networks. It has developed over the past decades as an open platform for innovation with low access barriers for end-users, providers of content, applications and services and providers of internet access services [35]. Access to the Internet is typically obtained by subscribing to an internet access service, a publicly available electronic communications service that provides access to the internet, and thereby connectivity to virtually all end points of the internet, irrespective of the network technology and terminal equipment used. [35]

The internet plays a vital role in our daily life as it is the medium through which services such as voice calls, web browsing, television and teleconferencing are offered to private, corporate and institutional customers. Nowadays, many of these services are considered an indispensable part of our lives.

There are three reasons to use the internet as starting point when discussing network virtualisation:

- Many networks in operation today are designed following the core design principles and technologies originally designed for the internet. The remainder of this section will survey these principles;

- Many networks today are part of the internet. Network virtualisation technologies are likely of the greatest relevance to these networks. Section 2.2 summarizes the challenges for network operators and section 2.3 provides an answer to these challenges by introducing network virtualization.

- The societal impact of implementing network virtualisation technology is likely to be the greatest when it is done on networks that are part of the internet. If network virtualisation is used to make private networks more flexible or efficient, this will most likely only lead to a one-sided competitive advantage for the owner of that network, whereas the potential societal impact of network virtualisation on the internet's networks may be much larger. Section 2 provides insights to the technological enablers while sections 3 and 4 build on these to indicate societal and economic impact.

2.1.1 Logical structure

The internet is a global collection of interconnected networks linking together billions of devices, which has as goal to carry information resources and services. The logical structure of the internet is summarized in Figure 6.

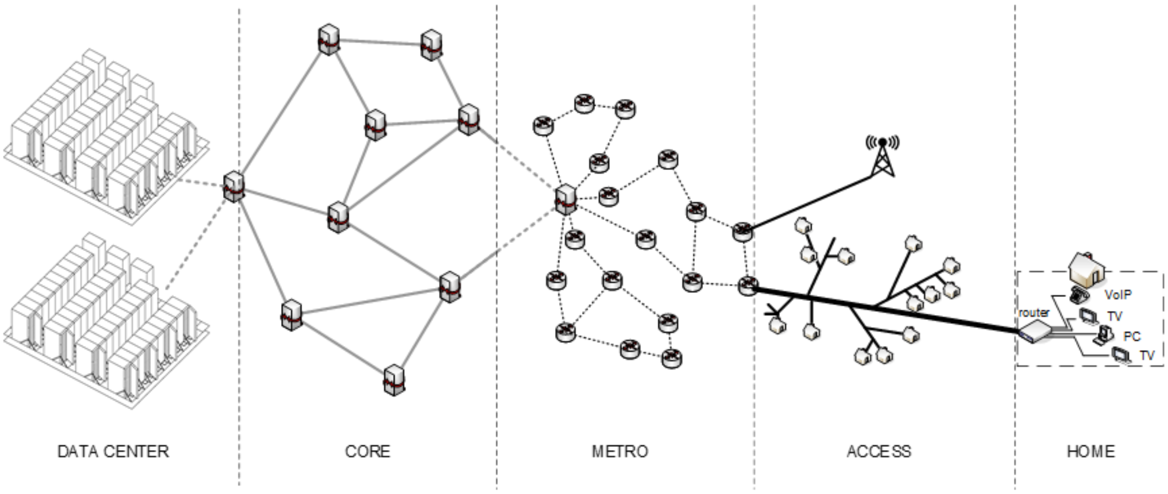


Figure 6. Home, access, metro (aggregation), core and data centre network

Home networks

Most end users, those who consume internet services, have a *home (or small company) network* deployed that consists of a limited number of desktop computers, VoIP handsets, TV sets, et cetera, which are interconnected via a wired or wireless local area network (LAN). The LAN is connected to the other networks of the internet via a router, which sits at the border of the home network and the access network.

Access and metro (aggregation) networks

The home network is connected by the *access and aggregation (or metro) network* to the backbone network. The *access network* is often referred to as the local loop or last mile as it spans only the last couple of kilometres between the provider's equipment and the subscriber. In the Netherlands, there are different types of access network infrastructures: copper cabling (which was originally deployed for the fixed telephone network), coax cabling (which was originally deployed for broadcast television), optical fibre (which is currently being deployed as the next-generation infrastructure for data) and wireless. In the Netherlands, wireless access is only used in limited cases as a substitute for fixed residential internet connections. Mobile internet access however is used heavily.

Access networks typically have a 'tree-like' structure, where the network aggregates more and more traffic the farther you move from the subscriber to the backbone. The *aggregation network* interconnects several access networks via a star, ring or meshed topology. They consist of tens, hundreds or even thousands of nodes typically interconnected by optical fibre, and aggregate all traffic from access networks towards core networks. Traditionally, in the aggregation network, circuit-switching technology has been used. As traffic often originates from an IP-enabled end-host and is packet-based, operators of aggregation networks typically at some point migrate to packet-switched technologies for this purpose (which gives opportunities for statistical multiplexing[3]).

Internet Service Providers (ISPs) such as KPN, Telecom Italia and Deutsche Telekom own both the access and aggregation networks and use these transport networks to offer X-play (e.g. triple play, quadruple play) services. In a triple play bundle, a subscriber obtains internet access as well as telephony and television service from the same operator. In a quad play bundle, mobile telephone is also included. In some cases, operators refer to what is called a *six pack* if the mobile subscription also includes mobile data and mobile television services.

In the Netherlands, there are several ISPs that operate without owning the access and/or aggregation networks. They make use of so-called *network access,* where they buy capacity on access and aggregation networks and subsequently run their own service over it. There are different forms of network access; a common type is *unbundled local loop*, which basically means that an operator is allowed to physically connect the provider end of a copper line (from the network of KPN) to their own equipment to provide DSL service to the customer. Other forms are *bitstream access* or VULA (Virtual Unbundled Loop Access), which provide similar service but virtually (i.e. without the physical access to the copper line). The ultimate form is called *wholesale access,* where the guest operator buys most of the components required to deliver service from the host network operator, or *white label*, where the host network provides the actual services, but they are branded differently.

Core networks

The *core network* (sometimes referred to as the backbone network) forms the core of the internet network to which the aggregation nodes are interconnected. It consists of about 40,000 autonomous systems[4] (AS) or domains. These networks transport the bulk of internet traffic, are based on optical transport technologies and consist of high bandwidth pipes responsible for transporting huge traffic volumes over large distances (e.g. a submarine optical cable crossing the Atlantic Ocean). These autonomous systems are interconnected via a large meshed topology and are structured in ‘tier levels’. Tier 1 ISPs participate in the internet solely via settlement free interconnection also known as settlement-free peering[5]. The typical characteristic of Tier 1 network is that they can reach every other network on the internet without paying internet transit[6] or paying settlements. Tier 2 networks peer with some networks, bus still purchase IP transit or pay settlements to reach at least a portion of the internet. Tier 3 networks solely purchase transit from other networks to participate in the internet. Note that it is difficult to determine whether a network is paying settlements as the business agreements are typically covered under a non-disclosure agreement (NDA).

Data centre networks

Whereas the network elements described above provide the connections, data centres contain the 'content' that can be accessed via the internet. Until around 2007, internet inter-AS traffic was dominated by ten to twelve large transit providers (tier 1 ISPs) interconnecting thousands of tier 2, tier 3, regional providers, consumer networks and content/hosting

[3] Statistical multiplexing and its benefits will be discussed in detail in section 2.1.2.

[4] A collection of connected Internet Protocol (IP) routing prefixes under the control of one or more network operators who presents a common, and expectedly consistent routing policy to the internet.

[5] Peering refers to two (or more) autonomous systems that interconnect directly with each other to exchange traffic.

[6] Internet transit is the service of allowing network traffic to cross a network.

companies. [24] Today, large over-the-top[7] service providers (OTT SP) such as Google and Microsoft operate their own data centres (e.g. as part of a content distribution network[8], CDN). These data centres are typically directly connected to several ISPs. It is advantageous for CDNs to peer with ISPs because the CDN does no longer need to purchase transit traffic (peering may not be settlement free). It also provides better throughput, higher reliability and lower network latency. At the same time, the ISP is able to provide its customers with good performance for a particular service, increasing its attractiveness for end users.

2.1.2 Fundamental design goals of the internet

The fundamental design goal of the internet was multiplexed (see below for an explanation) utilization of existing interconnected networks. This means (1) that several senders of data (e.g. user A browsing a web page and user B watching IPTV) use the same communication channel and (2) that the communication channel is realized as much as possible via existing networks. There are two fundamental challenges to this goal: (1) shared use of a single communication channel and (2) the interconnection of existing networks.

Statistical multiplexing[9]

The first challenge was conquered by using *statistical multiplexing* or *packet switching technology*.

The advantage can best be explained by first considering circuit-switched technologies. Circuit-switched technologies pre-allocate and reserve circuits regardless of the effective demand. A circuit in a link is implemented with either frequency-division multiplexing (FDM) or time-division multiplexing (TDM). Proponents of packet switching argue that circuit switching is wasteful because the dedicated circuits are idle during silent periods. This is clarified in Figure 7.

If user A wants to communicate with user C, the network establishes a dedicated end-to-end circuit between the two hosts. Thus, in order for user A to send messages to user B, the network must first reserve one circuit on the link between the two circuit switching nodes. If the link has *n* circuits, each end-to-end circuit over a link gets the fraction 1/*n* of the link's bandwidth for the duration of the circuit (Figure 7 assumes *n* = 1). However, if several circuits arrive at a common switching node to use the same outgoing link, no benefit can rise from sharing the common path. For example, if user A uses a circuit-switched network to remotely access photographs stores on the computer of user C, the user sets up a connection, requests an image, contemplates the image, and then requests a new image. Network resources are wasted during the contemplation periods. For example, if user B tries to communicate with user C or D, it will receive a busy signal from the network. In order to guarantee that both circuits can follow the same path, *n* must be equal to the *sum* of the required circuits by the incoming links.

[7] Over-the-top refers to delivery of content (audio, video and other media) over the Internet without the ISP being in control of the distribution of the content.

[8] A content delivery network (CDN) is an overlay network of web caches that are geographically spread across the world in different data centres. Non networked CDNs such as Google and Akamai typically place their servers in other ASs or ISPs. Large content providers such as google (e.g. YouTube) may also interconnect the servers with their own global backbone network.

[9] Statistical multiplexing and network virtualisation are two distinct subjects. Where statistical multiplexing allows several senders to share a single link, network virtualisation allows diverse network architectures to co-habit on a network of physical resources (a combination of link and node resources). Network virtualization is covered in more depth in paragraph 2.3.

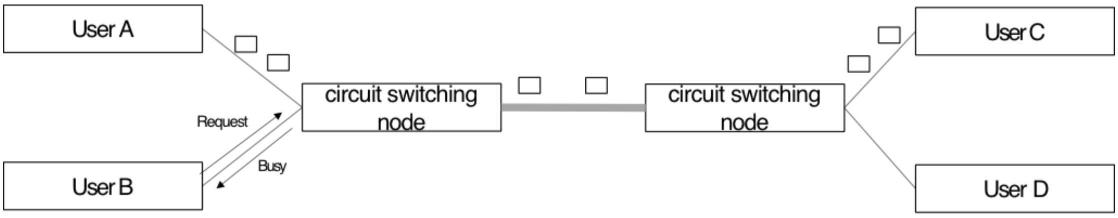


Figure 7. Circuit switching

Packet switched technologies on the other hand, can share the same link by only using bandwidth if packets are effectively received. Many senders can send data over the same network at the same time, effectively sharing the resources in the network. Contrary to *circuit switching*, there is no state established ahead of time and there are only few guarantees regarding the level of service that the network provides (best effort). This is illustrated in Figure 8.

Suppose hosts A is sending packages to host C and host B is sending packages to hosts C and D. The packet switches direct these packets to the correct user. If there is congestion at a link, the packets queue in the link’s output buffer before they can be transmitted over the link. As shown in Figure 8, the packets do not follow any periodic ordering. The ordering is random or statistical because packets are sent whenever they happen to be present at the link. For this reason, packet switching is said to employ statistical multiplexing. An advantage of statistical multiplexing of the links and the network, is that the sender never gets a busy signal (in contrast to circuit switching). Disadvantages of packet switching are the variable delay and the potential for dropped data packets.

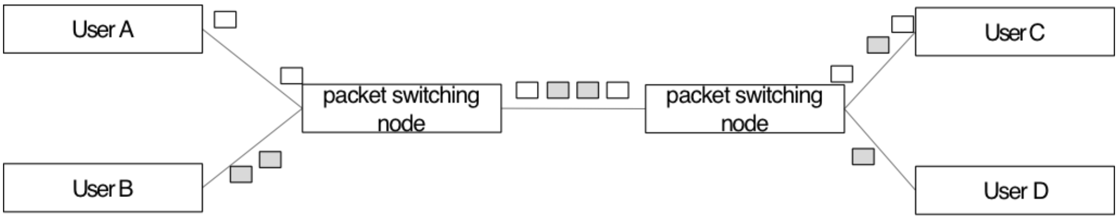


Figure 8. Statistical multiplexing as possible in packet switching

The narrow waist

The second design challenge was solved by designing the so called narrow waist which provides the network logic to interconnect the underlying physical networks and separates the application logic from the network logic. From a technological point of view, the internet is structured according to the layered TCP/IP model, containing five layers depicted in Figure 9.

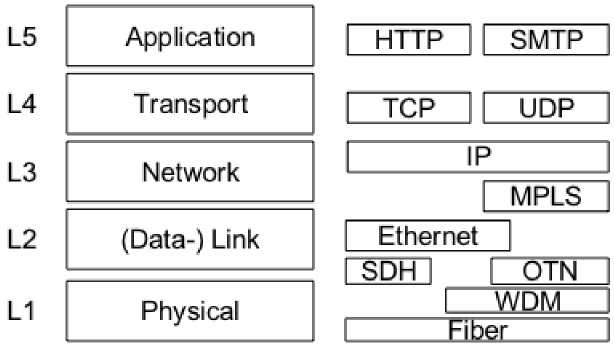


Figure 9. Technologies within the TCP/IP layering

In this (theoretical) model, every lower layer in the model provides service to a higher layer. At the centre layer (the *network layer*), is an interconnection protocol, implemented by the Internet Protocol (IP). To connect to the internet, a device must implement the IP stack. The network layer guarantees end to end connection-less connectivity[10]. Thus, if a host has an IP address, then the network layer provides the guarantee that a packet with that host destination address should reach the destination with the corresponding address (with best effort). This core function of IP is reached by providing the following services to higher layers: (1) connection-less connectivity between end-hosts (packet-based messaging), (2) node addressing and address aggregation of end-hosts and intermediate nodes, and (3) efficient message forwarding and path determination (routing) between source and destination nodes via intermediate gateways or routers.

On top of the network layer sits the *transport layer*. The transport layer includes protocols like Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). Transport layer protocols provide various guarantees to the *application layer* including port numbers for addressing different functions at the source and destination of the datagram, checksums for data integrity, reliable transmission, flow control, congestion control, etc. The application layer includes many protocols that various internet applications use such as the hypertext transfer protocol (HTTP) and the Simple Mail Transfer Protocol (SMTP) allowing typical internet services and applications such as web-browsing and e-mail.

Below the network layer, the (data-) link layer provides point to point connectivity between individual nodes, or connectivity on a LAN. Ethernet is an example of a link layer protocol. Below the datalink layer, the physical layer ensures transmission of the data over a given medium via protocols such as Synchronous Digital Hierarchy (SDH).

1.1.1. Network planes

Communication networks do not only transport end-user data, but also need to exchange control-related data and implement related functionality to guarantee that the network operates as designed. The key functionalities of a network are typically divided into three planes: the data plane, the control plane and the management plane.

[10] A connection-oriented protocol is one where a logical connection is first established between devices prior to data being sent. In a connectionless protocol, data is just sent without a prior connection being established between devices and the source does not attempt to monitor whether data is delivered to the destination.

The data plane

The data plane contains all functionality that relates to the transmission of end-user data (payload) in the network. The data plane carries out the commands of the control plane.

The data plane is responsible for the transmission and reception of data packets, including packet buffering, packet scheduling, header modification and forwarding at individual nodes to send the data to the next node. It consists of a number of ports. The correct route is determined from a so-called FIB (Forwarding Information Base). Next to the FIB, it consists of a number of ports which are used for the reception and transmission of packets.

The control plane

The control plane contains all functionality that is responsible for the correct configuration of the data plane. The control plane is responsible for the exchange of status information, such as host reachability, with neighbours (discovery function). It also decides how data must be forwarded in the network (routing function) and performs the reservation (path setup) and release (path breakdown) of required resources.

The control plane is the brain of the router and consists of routing protocols, such as OSPF (open shortest path first), BGP (boarder gateway protocol), IS-IS (intermediate system to intermediate system) and several other protocols such as IGMP (Internet Group Management Protocol), ICMP (Internet Control Message Protocol) and so on. The control plane also contains the RIB (Routing Information Base). This is the routing table where all IP routing information is stored. The RIB is updated when a routing protocol learns a new route or when a destination becomes unreachable. The RIB may also contain routes which are added by an administrator (static routes) as well as back-up routes to the same destination. Between the control and data plane, a communication channel (or interface) is used to insert routes from the RIB into the data plane’s FIB (Forwarding Information Base).

The management plane

Some management related operations are not considered as control functionality. The management plane provides the interface to the network operator for performing such management operations, and allows further configuration and monitoring.

For further clarification, the control and data planes of a router are illustrated in Figure 10.

In commercial routers, the control plane typically runs on low-end CPU (central processing unit). In contrast, the data plane uses special-purpose high speed lookup memory (such as Ternary Content Addressable Memory, TCAM) to store entries. As such processing of packets is slower in the control than in the data plane. The control and data plane are tightly integrated in commercial routers. This approach has been highly successful as illustrated by the success of the internet. It has however two disadvantages. First, the communication channel between the data and control plane in commercial routers is a proprietary and closed implementation. As such the evolution of both data and control plane are closely tied together. Second, special-purpose hardware such as TCAMs are costly and have high power consumption.

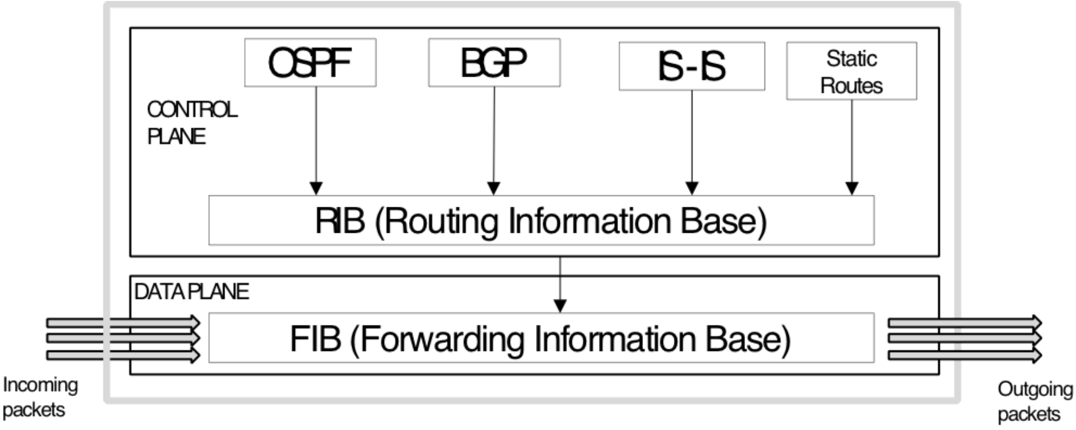


Figure 10. Basic router design

2.1.3 The internet today

The internet has been stunningly successful in doing what it was designed to do: enabling data communication. As such it has shaped the way we access and exchange information in the modern world [13]. Today, the internet architecture supports a multitude of applications and it is able to run over a wide variety of physical networks. The internet infrastructure is being continually upgraded to cope with growing demands in terms of, among others, performance, reliability and scalability. As a proxy to indicate this process, the number of IETF[11] Request for Comments (RFC) publications is visualized in Figure 11. An RFC is a formal proposal describing an addition to the suite of technologies considered to be an integral part of the internet itself.

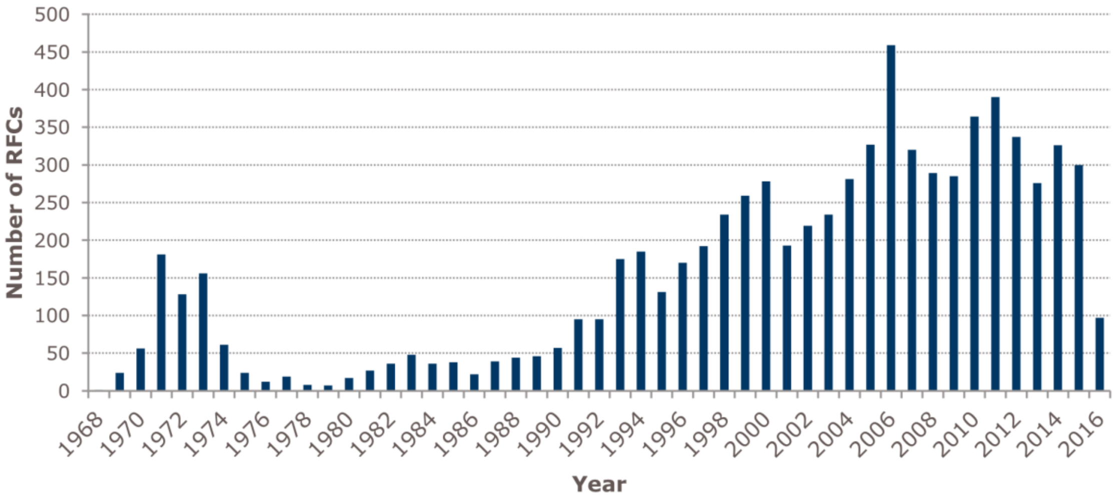


Figure 11. IETF RFC publication rate per year

The popularity of the internet makes further growth difficult as it has made radical changes and introduction of new network architectures nearly impossible. For example, new types of services (e.g. Internet-of-Things related services) may pose requirements that are

[11] The mission of the Internet Engineering Task Force (IETF) is to make the Internet work better by producing high quality, relevant technical documents that influence the way people design, use, and manage the Internet.

challenging to meet with the current internet network architecture (e.g. the IPv4 address space may not be adequate to provide end-to-end connectivity for all connected devices). Due to its decentralised, multi-provider nature, adopting a new architecture or modification of the existing one requires consensus among competing stakeholders (see, for example, the slow adoption of the IPv6 protocol and the popularity of network address translation, NAT, as an alternative). As a result, alterations to the internet architecture have become restricted to 'simple' incremental updates (e.g. NAT) breaking the initial design principles (e.g. end-to-end principle). This approach of patching the internet architecture with small incremental updates has been in general successful as manifested by the number of users and the multibillion euro industry around it. However, at the same time, these incremental updates increase the complexity of the internet and prove challenging for network operators.

2.2 The challenges for network operators

This section provides an overview of the challenges that network operators face today. We focus on data centre operators and internet service providers but would like to note that transit network operators, internet exchange points, et cetera face very similar challenges.

2.2.1 Data centre operators

The Internet as a system of interconnected computer networks provides the end-user with access to data which can be hosted on a server located on the other side of the planet. For example, a user surfing to a website (e.g. www.netflix.com) may be served from a server hosted in the United States. To provide such a service, a combination of resources is required:

- **Server resources**, which consists of two separate types of resources:
  - *Storage resources*: to store the webpage file (e.g. an HTML file) and video content (e.g. in SD, HD and Ultra HD video quality)
  - *Compute resources:* to build the website from the stored webpage files, to playout the video files, to interpret and react to user prompts, et cetera.
- **Network resources:** to provide a communication channel between the server and the consumer (the combination of network nodes and links)

Content, like Netflix's videos, that is provided via the Internet without the direct involvement of ISPs is referred to as over-the-top content. Its providers are often referred to as over-the-top service providers (OTT SPs). They use the Internet to offer a heterogeneous set of internet services to the end user including online gaming, audio streaming (e.g. Spotify), video streaming (e.g. Netflix), cloud storage (e.g. Dropbox) and cloud computing (Google's App Engine), etc.

As the popularity of these services grows, so does the amount of server and network resources that is required to provide them. Today, large OTT SPs such as Netflix, Google and Facebook are responsible for the majority of all network traffic. Similarly, the server resources resemble a warehouse full of computers [6]. These new large datacentres (DCs) are quite different from traditional hosting facilities of earlier times and cannot be viewed simply as a collection of co-located servers. They differ significantly from traditional datacentres: they belong to a single organization, use a relatively homogeneous hardware and system software platform, and share a common management layer. Often, much of the application, middleware, and system software is built in-house compared to the predominance of third-party software running in conventional datacentres. Most importantly, warehouse-scale computers run a smaller number of very large applications (or Internet services), and the common resource management infrastructure allows significant deployment flexibility [6]. Large portions of the hardware and software resources in these facilities must work in concert to efficiently deliver good levels of internet service performance.

The services provided by OTT SPs rely on the data centre LAN network and the WAN network (connecting data centres) to tie the server resources together. To reach these benefits, DC operators have to overcome a number of challenges:

Inflexibility in the network

A data centre needs to support a heterogeneous set of services. These services put stringent requirements on the data centre network (e.g. QoS, latency, fault tolerance). New services can also have unpredictable demands of the network. Adding new features to conventional network equipment is considered as a lengthy process, leading to a long lead time from the initial service conception to the realization.

High network cost

For large scale data centres, virtualisation technology has driven down the cost of storage and computing resources. In contrast, due to the bundling of hardware and software (in many cases including unnecessary features), conventional network equipment remains expensive.

2.2.2 Internet service providers (ISPs)

The European telecommunications sector has undergone drastic changes during the last decades. From the late nineties, privatization and liberalization of telecommunication networks was initiated. The legacy copper connections which provided incumbents with direct access to customers' physical location positioned incumbents favourably to enjoy from the digitalization of the European industry and the increasing number and demand of consumers.

Soon incumbents started to diversify their products and digital services to include services such as mobile voice calls, SMS, broadband internet and digital television in addition to the original fixed line voice call service. A number of challengers to the incumbents appeared on the market and today a consolidation via mergers and acquisitions is taking place possibly leading to five major European ISPs: Orange, Telecom Italia, Vodafone, Deutsche Telekom and Telefonica. The current business model of these providers can best be described as a combination of broadband (mobile) internet access (a "dumb pipe" through which data can be transported) and a limited set of value added services that are offered via the same transport network (e.g. digital TV). This has resulted in a highly profitable business (Figure 12) which gives, at first sight, the impression to be quite future proof, in the face of developments such as cloud computing and connected devices.

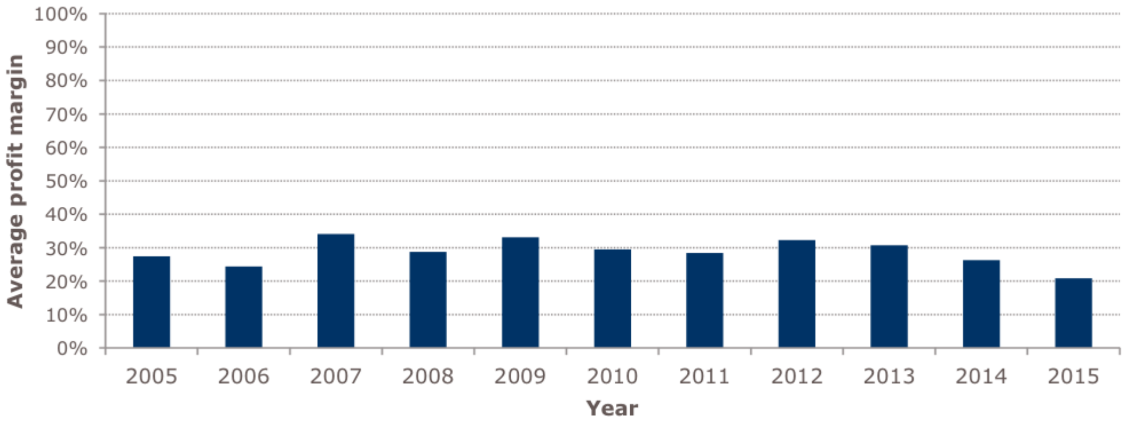


Figure 12. Average profit margin[12]of five major European ISPs (2005-2015)

That business model is however challenged in several ways. The main challenges are: (1) a shift in revenue stream sources, (2) traffic growth, (3) national and European Union service expectations and regulation (4) networks that are inflexible towards service innovation, complex and costly to operate.

The shift in *revenue stream sources* and the *exponential growth in traffic* is driven by the ongoing digitalization[13] of the value network of ISPs and leads to a shift from paying for connectivity towards paying for content. The abundance of services enabled by the internet is changing the economics as well as the origin, type of and amount of traffic flowing over ISP networks. This may lead to a shift in where revenue is accrued along the value chain. Broadband internet subscriptions offered by ISPs mostly operate on a pricing model that provides customers with flat rate access to (un)limited internet data. The ISP as such has to invest in increasing its network capacity (e.g. in the local loop) while the OTTs are the primary recipient, being able to generate revenue from their services offered over the ISP's network[14]. In addition, several OTTs provide services which directly compete with the value-added services offered by ISPs, further pressuring ISP revenues[15].

Europe's telecommunications infrastructure is regulated by national and European bodies that monitor its development. Some services and the network as such must satisfy requirements from regulators. Examples of *regulation* are service to rural areas, quality of service requirements, universal service obligations, roaming rules, etc. In addition, both the European commission and national bodies have put forward plans that state *increased service specifications*. Broadband Europe for example stipulates access to 30 Mbps connectivity to every European and wants half of the households to have the possibility to subscribe to a 100 Mbps connection by 2020. To reach this objective, many ISPs will need

[12] The profit margin was calculated by dividing earnings before interest, taxes, depreciation and amortization (EBITDA) by the total revenue.

[13] Digitalization refers to the use of digital technologies to change a business model and provide new revenue and value-producing opportunities.

[14] ISPs benefit indirectly from the success of OTT providers as the OTT has to pay a transit fee to reach the ISPs customers. OTTs also invest in deploying own network capacity.

[15] One example is cord cutting where viewers cancel their TV subscription with ISPs in favour of competing services from OTT service providers.

to upgrade their legacy connections to high capacity optical fibre requiring a considerable upgrade expense while the return on investment is deemed inconclusive.

In reaction to these challenges, several ISPs have chosen to focus on one specific market segment, such as the enterprise market while others have chosen to diversify their products and digital services to compete with digital platforms. ISPs are favourably placed to do so as they are able to integrate services closely into the network to provide a high-quality of service. During this transition, ISPs face a number of challenges:

Inflexibility

The ISP's network infrastructure has to support new services, to keep up with market developments. Today it is not possible to quickly add extra functionality to existing network devices. The service release cycle is long as, first, standards development organizations need to agree on a standard during the standardization process and next, vendors need to approve and incorporate new solutions in operating networks. A thorough evaluation of the solution has its merits as network failures should be prevented and network uptime should be maximized but it may also lead to unnecessary delays. These delays may force network operators to rely on old legacy equipment that is not capable of providing the required support for emerging services and results in the loss of business opportunities.

High complexity

ISP's network operators have to cope with an abundance of *legacy technologies and systems* that reside in their networks. In addition, current telecommunication networks are characterized by large deployments of middle boxes (Figure 13), providing L4-L7 network services. These middle boxes offer valuable benefits, such as improved security (e.g. firewalls and intrusion detection systems), reduced bandwidth costs (e.g. WAN optimizers) and improved performance (e.g. proxies).

These have complex and specialized processing, variations in management tools across devices and vendors, and imply a need to consider policy interactions between an appliance and other network infrastructure. ISPs require trained staff to 'manually' configure and reconfigure devices. Due to the high level of complexity, manual configuration is error prone and may result in misconfigurations resulting in service disruptions.

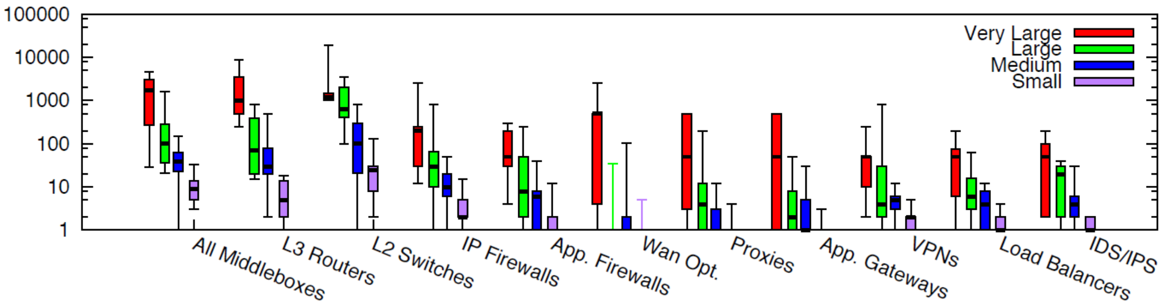


Figure 13. Box plot of middle box deployments for small (fewer than 1k hosts), medium (1k-10k hosts), large (10k-100k hosts), and very large (more than 100k hosts) enterprise networks. Y-axis is in log scale [39].

High cost

Today, network devices implement the majority of the IETF RFCs in their control plane (see Figure 11). This clearly increases the cost of software development[16]. Furthermore, vendors code their implementation of a standard in a closed environment, in contrast to open source software (OSS) development. The vendor implementations of a standard should allow for heterogeneous devices from multiple vendors to function in concert. In reality, vendors enhance these standards to differentiate themselves from competing vendors. This however often results in devices of one vendor not operating smoothly with products from other vendors.

2.2.3 Consequences of these challenges

We argued in the previous section that network operators are challenged by the increase in complexity of the internet architecture. Because of this, network operators have to deal with increasingly complex and inflexible network technology which comes at a high cost and has low potential for service innovation and revenue growth.

Further evolution along this line is not wanted. Network operators will have to deal with higher capital and operational expenditures (CAPEX/OPEX) at a time when average revenue per user (ARPU) is decreasing. Progress down this evolutionary path is bound to incur an increasing rate of costs in the future. Meanwhile, operations staff have to maintain many types of network elements and follow the evolution of dozens of innovative technologies. Second, as higher CAPEX / OPEX forces some operators to refrain from investing further, those who do invest face long time-to market periods as it is challenging to add new features. Interested operators must push a whole industry to standardize these new features, and then wait for vendors to actually implement them. This process may assure quality standards development, but requires a lot of effort, and prevents operators who are willing to endeavour first into new domains from doing so quickly [34].

2.3 Network virtualisation: a way out?

Network virtualisation, software-defined networking and network function virtualisation are three distinct but related proposals that have as common goal to break out of this downward spiral.

Network virtualisation allows multiple heterogeneous network architectures to cohabit on a shared physical substrate [13]. By doing so, network virtualisation enables innovation and a way out of the downward spiral described in 2.2.3.

To realize a network that allows for diverse network architectures to be run on a shared physical substrate, the separation of policy from mechanisms is promoted[17].

[16] Network vendors implement an extensive set of features while a network operator may only require a subset of these features. Instead of paying for the development of only those features required network operators end up paying for all available features.

[17] Policy refers to the specification of the manner in which a set of resources are managed. Mechanisms, on the other hand, refers to those parts of a system implementation that control the authorization of operations and the allocation of resources (the means by which policies are implemented). Policy/mechanisms separation is the segregation of the entities that dictate resource management strategies from entities that implement the low-level tactics of resource management. This idea was introduced by Per Brinch Hansen in his work on operating systems.

When the separation of policy from mechanisms is promoted, the role of the traditional ISPs can be divided into two: infrastructure providers (who manage and operate the physical infrastructure) and service providers (who create virtual networks by aggregating resources from multiple infrastructure providers and offer end-to-end services to the end users (Figure 14). Such an environment will foster deployment of multiple coexisting heterogeneous network architectures that are not bound by the inherent limitations found in the existing internet [13].

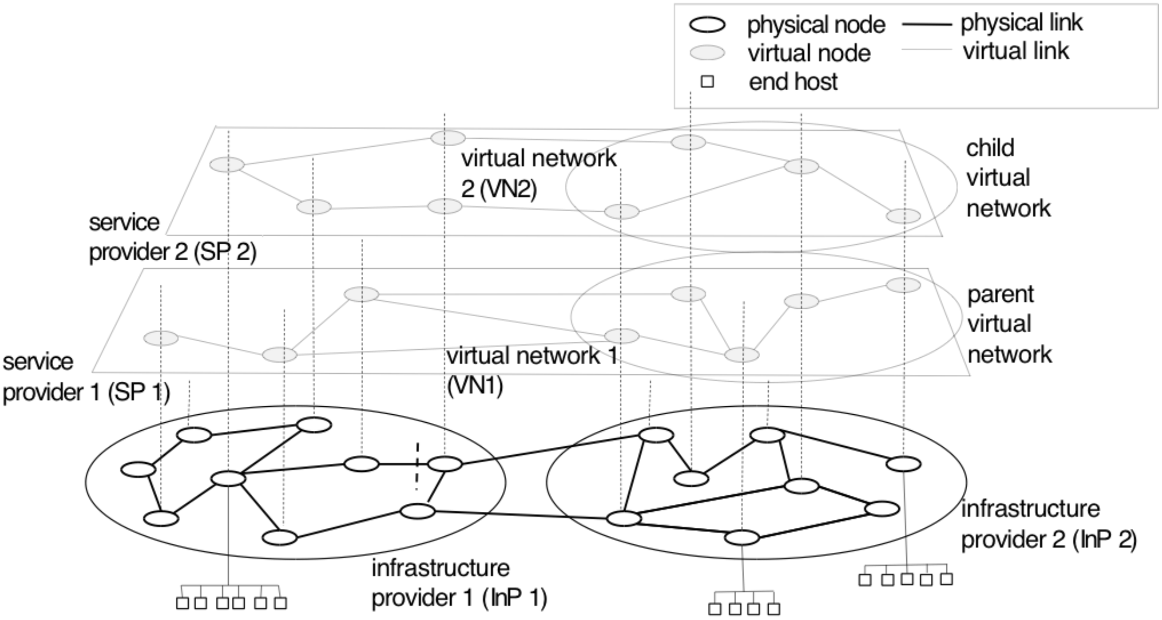


Figure 14. Network virtualisation architecture

As will be discussed in more detail below, network virtualisation provides flexibility, promotes diversity and promises security and increased manageability.

2.3.1 Network virtualisation and programmability

To make network virtualisation possible, programmability of the network elements is of utmost importance [13]. Only through programmable network elements it will be possible for the service providers to implemented customized protocols and deploy diverse services. Hence, the design decisions: "*how much programmability should be allowed?*" and "*how it should be exposed*" must get satisfactory answers. The level of programmability refers to the level of detail at which programmability is allowed. Examples are at the level of individual packets or at the level of flows of packets. More detail allows for more flexibility at the cost of a more complex programming model. The exposure of programmability refers to who should be allowed to program the network. One extreme is that each user should be allowed to execute any new code while on the other end of the spectrum only a small set of users may only be allowed to call functions that are already available.

- **Programmable control plane:** The software-defined networking paradigm is one possible answer to these questions. Software-defined networking enables network operators to configure the control of their networks through their own custom software. We refer the interested reader to section 2.3.2 for an introduction to software defined networking.

- **Programmable data plane:** Network function virtualisation pushes the programmability of the network even further by making it possible to code data plane behaviour in software, enabling it to run on general purpose server hardware rather than on expensive vendor-controlled hardware platforms. We refer the interested reader to section 2.3.3 for an introduction to network function virtualisation.

NFV and SDN are two closely related technologies they are not necessarily dependent on each other but they can benefit from each other. While NFV goals can be achieved without the separation of data and control plane or the centralization of network control, usage of SDN can simplify the configuration of a virtual network function. NFV on the other hand could benefit SDN by providing the infrastructure upon which SDN software can run (e.g. an SDN controller could be hosted in a virtual machine that runs on general purpose hardware).

2.3.2 Software-defined networking: programmable control plane

The goal of software-defined networking is to allow network operators to reduce the complexity of network configuration. To allow so, the network's configuration is centralized in a logically centralized controller. SDN offers, via a logically centralized controller, two things. First, a network-wide view of both topology and traffic which allows network operators to define and satisfy network level objectives (e.g. load balancing, security, etc.). Second, direct control of the data plane rather than indirect configuration of each individual device.

This results in a separation of concerns (between the control plane and data plane). The router hardware, which is specialized to forward traffic at very high rates, should forward packets and collect measures such as traffic statistics and topology information. The logically centralized controller on the other hand, focuses on the computation of routes (routing). This is fundamentally different from conventional router designs where routing has operated as a distributed computation of forwarding tables.

The promise of SDN is to reduce the complexity of network configuration which will hopefully lead to better network configuration as it is easier to coordinate and reason about the behaviour among a network of devices (*reduced complexity and cost)*. SDN provides the additional benefit that the control and data plan are able to evolve independently (*increased flexibility)*. As the control plane is no more than a software program written in a high level language, it becomes easier to introduce new features and as such to spur innovation. The data plane on the other hand is typically programmable hardware with a longer development cycle. In SDN, the data plane functionality could run on commodity hardware which can be purchased at a lower price (*reduced cost*).

Architectural concepts of SDN

The term SDN was originally coined to represent the ideas and work around OpenFlow at Stanford University [23][18]. As originally defined, SDN refers to a network architecture where the forwarding state in the data plane is managed by a remote control plane decoupled from the former (Figure 15A and B).

[18] SDN does not require OpenFlow as the southbound interface.

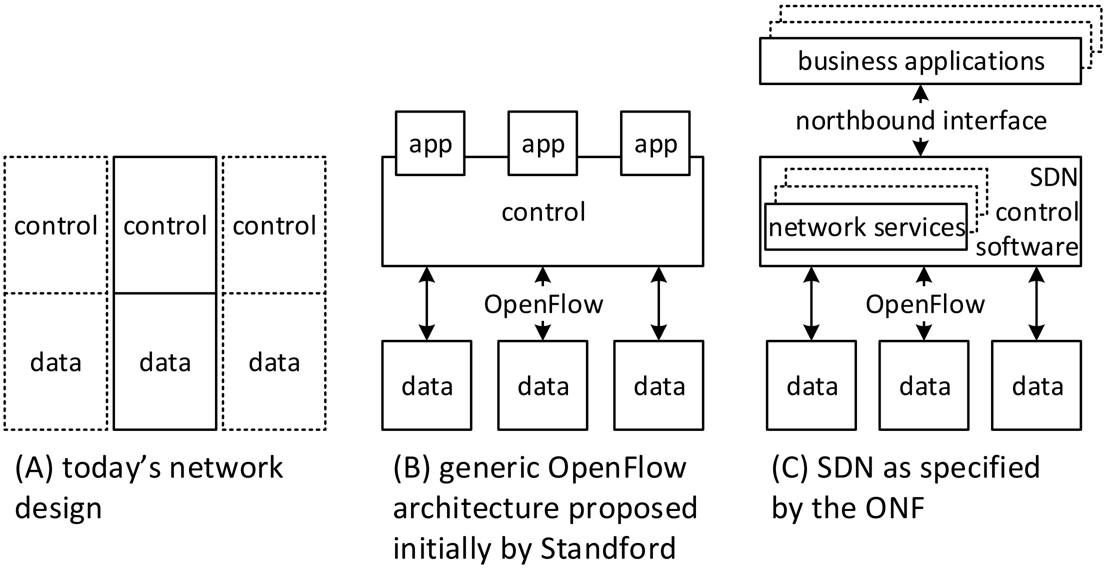


Figure 15. Conventional, Stanford and Open Networking Foundation (ONF) network design

By removing the control functionality from network devices, these latter become simple (packet) forwarding elements. The control functionality (logic) is moved to a (logical) centralized external entity, the so-called SDN controller or network operating system (NOS). The NOS is a software platform that (typically) runs on commodity server technology and provides the essential resources and abstractions to facilitate the programming or forwarding devices based on a logically centralized, abstract network view [23]. Its purpose is therefore similar to that of a traditional operating system.

In addition, the network is programmable through software applications. These software applications run on top of the NOS. The design proposed by ONF (Figure 15C) makes this explicit by placing applications in a separate entity (i.e. business applications). These business applications can communicate with the controller through the northbound interface of the controller. The controller provides an abstracted view of the network to these business applications which can in turn use that view to provide appropriate instructions to the control plane layer to perform specific actions in the data plane. The availability of open interfaces, in contrast to proprietary interfaces in conventional commercial network equipment, is considered as a major differentiator for SDN network equipment.

A final characteristic of SDN is that, forwarding-decisions are made on a flow-basis, instead of destination-basis. A (packet-)flow is broadly defined as a sequence of packet field values which can be uniquely identified by parameters such as: (1) source IP address, (2) destination IP address, (3) source port, (4) destination port, and (5) layer 4 protocols (TCP/UDP). For instance, when a user surfs to a website, a new flow will be created with the following parameters: (1) the transport protocol: 6 (TCP) or 17 (UTP), (2) source port, e.g. 1234, (3) destination port, e.g. 80, (4) source IP, e.g. 157.193.240.2, and (5) destination IP: the IP address of the website.

Packets from one flow can be handled differently from others, by means of separate queues/actions. Therefore, using flow parameters, packets of different flows can be distinguished to apply different actions (such as traffic shaping). In the SDN/OpenFlow context, a flow is a sequence of packets between a source and a destination. The packet field values that identify a flow act as a match (filter) criterion to which a set of actions (instructions) can be applied. All packets that are part of a particular flow receive identical service policies at the forwarding devices. By using the flow abstraction, the behaviour of different types of network devices (e.g. routers, switches, firewalls and middle boxes) can be mimicked.

2.3.3 Network function virtualisation: programmable data plane

NFV decouples network functions from dedicated hardware to allow these network functions to be hosted on a virtualized environment. Network functions[19] which are provided through software virtualisation techniques are referred to as virtualized network functions (VNFs). The main target of NFV are the network services that are now being carried out by router, firewalls, intrusion detection systems, load balancers, etc. (middle boxes).

By decoupling the NF from dedicated hardware NFV aims to achieve six goals [19]:

- The *first goal* is to reduce cost compared with dedicated hardware implementations. This can be achieved by using standard hardware (i.e. general purpose servers and storage devices) to provide network functions (NFs) through software virtualisation techniques. Sharing of hardware and reducing the number of different hardware architectures in a network also contributes to this objective.

- The *second goal* is to improve flexibility in assigning virtual network functions to general purpose hardware. By hosting network functions on virtual machines (VMs), it will become possible to add capacity through software. This aids scalability and largely decouples functionality from location, which allows software to be located at the most appropriate places. This facilitates resource sharing, enables time of day reuse and enhances resiliency.

- The *third goal* is rapid service innovation through software-based service deployment. The time-to-market can be improved as the evolution of (software based) network functions is no longer tied to specialized hardware.

- The *fourth goal* is to improve operational efficiencies resulting from common automation and operating procedures. As capacity can be scheduled more flexible, network operators will be able to respond in a more agile manner to changing business goals and network service demands.

- The *fifth goal is to* reduce power usage achieved by mitigating workloads and powering down unused hardware. The *final goal* is to define standardized and open interfaces between virtualized network functions and the infrastructure and associated management entities so that decoupled elements can be provided by different elements.

Architectural concepts of NFV

The standards development organization European Telecommunications Standards Institute (ETSI) provides a high-level NFV framework to enable dynamic construction and management of VNF instances and the relationships between them regarding data, control, management, dependencies and other attributes. This framework is illustrated in Figure 16 and identifies three main components of an NFV network. The first component includes a

[19] The term network function is not defined by the IETF (the standards development organization governing most NFV initiatives). It does refer to the specific operations that are undertaken by middleboxes on a packet (or flow) between the transmitter and the ultimate receiver. These are today typically done in specialized hardware.

diverse set of virtualized network functions (VNFs). These are the software implementations of a network functions implemented in such a way that they can run over the NFV Infrastructure (NFVI). The NFVI consists of the diversity of physical resources as well as the software tools (hypervisors) that enable virtualisation of the compute, storage and network resources (virtualisation layer). The NFVI resources are offered as virtual compute, storage and network resources to support the execution of VNFs. The NFV management and orchestration component focuses on all virtualisation-specific management tasks necessary in the NFV framework to cover the orchestration and lifecycle management of VNFs.

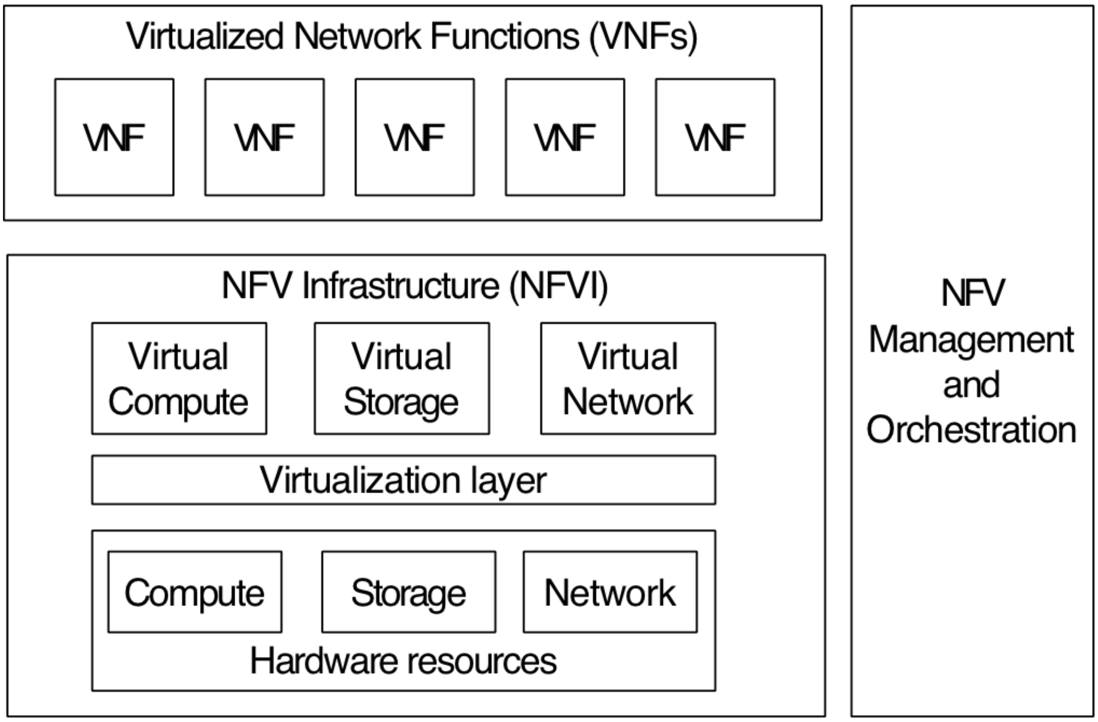


Figure 16. High-level NFV framework, adopted (source [19])

From the perspective of a telecommunications network operator who wishes to deploy a network-based service (e.g. mobile internet), the network connectivity between VNFs is important. A VNF forwarding graph (VNF-FG) defines the sequence of NFs that packets traverse for the case where network connectivity does matter. VNF FG are the analogue of connecting physical appliances via cables, in other words a VNF FG provides the logical connectivity between VNFs [20]. The decomposition of the service into NFs is referred to as service decomposition. By decomposing a service into elementary NFs, a number of benefits can be realized. First, re-usable elementary blocks are developed. Second, new and more complex services can be realized from these elementary blocks and third, the detailed implementations of these NFs can be abstracted. Figure 17 depicts an example service decomposition. The service graph is decomposed into three NFs (NF1, NF2 and NF3), NF2 is decomposed to NF4 and NF5, etc.

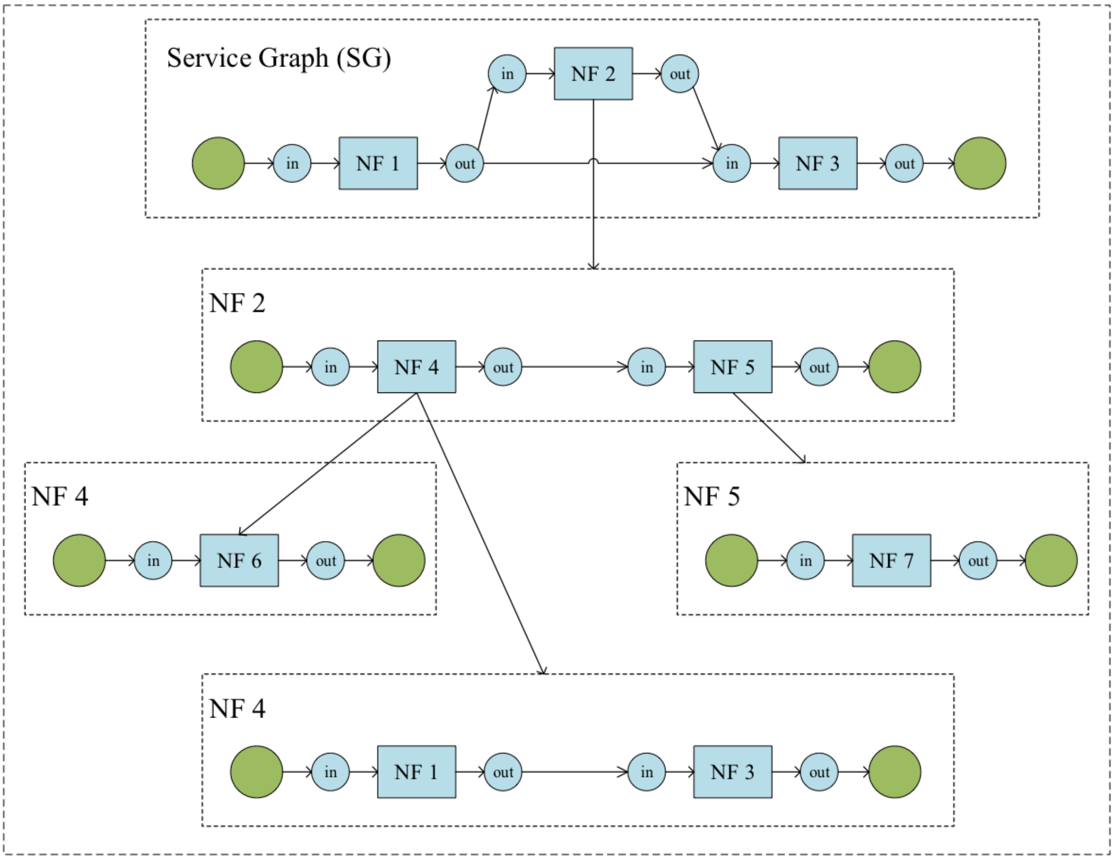


Figure 17. Example of service decomposition process

2.4 Overview

The internet has developed over the last decade as an open platform for innovation with low access barriers for end-users, providers of content, applications and services and providers of internet access services. The internet's own success has however made it increasingly difficult to introduce truly innovative ideas. As a consequence, network operators face a number of challenges such as networks that are expensive and complex to build and operate as well as a long time between service inception and realization.

Network virtualization is considered as a way out. Network virtualisation allows multiple heterogeneous network architectures to cohabit on a shared physical substrate. Software-defined networking and network function virtualization are considered as two enablers to realize network virtualization. Software-defined networking proposes the separation of the data from the control plane via a logically centralized control plane which communicates with the data plane via an open interface. Network function virtualization proposes the 'softwarisation' of middleboxes such as deep packet inspectors, network address translators, etc. to allow these network functions to run on virtual instances of standardized hardware.

Network providers can use network virtualisation to reduce operational complexity, lower expenditures and increase the rate of innovation. The next section discusses how different types of network operators react to network virtualization.

3 Implementation scenarios

In this chapter, we will briefly discuss the different types of networks that can be distinguished with respect to SDN and NFV implementation. We will then discuss the drivers for SDN and NFV adoption in each of these networks. Depending on these drivers, we then distinguish different scenarios for implementation of SDN and NFV in the networks.

3.1 The different playgrounds for SDN and NFV

The networks that are in operation today are highly heterogeneous from the perspective of the applications running over them, all the way down to the physical infrastructure that actually transports the data. On most networks however, data is transported using the *internet Protocol* (IP), and these networks are usually also connected to other IP networks, which allows them to connect to all other networks on the internet. We will therefore take internet-connected IP networks as a starting point for our analysis, and then discuss specific different types of networks wherever appropriate.

3.1.1 Network hierarchy on the internet

In the context of the internet, individual *logical* networks are referred to as *autonomous systems*, or *ASes*. An autonomous system is defined as a set of IP addresses[20] that are under the control of a single network operator. Note that there is not always a one-to-one mapping between a logical and a physical network, although in practice, a physical network will usually only be part of one or a handful of ASes. Some networks can be considered an autonomous system, but are not identified as such on the internet as they are completely private (or only have limited connectivity to the internet).

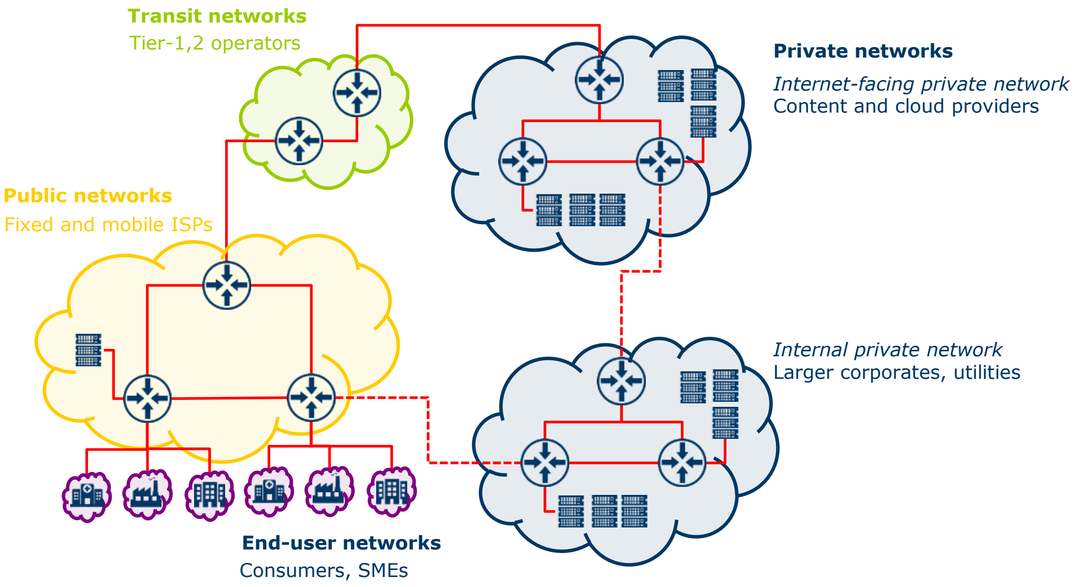


Figure 6 Different types of networks for which SDN and NFV are of relevance

[20] More precisely, an AS advertises a set of so-called 'prefix'. If an IP address starts with a prefix advertised by an edge router of an AS, the router will be able to deliver traffic to the host at that address.

The different ASes connected to the internet can fulfil different roles with respect to the functioning of the internet. The role fulfilled by an AS is highly relevant with respect to the types of SDN and NFV implementations that are useful. Figure 6 gives an overview of different types of ASes and their relative position to each other.

3.1.2 Transit networks and internet exchange points

For a host inside an AS to be reachable from hosts in other ASes, there needs to be a *route* between the hosts that data packets can follow to reach their destination. The route followed by traffic between two hosts is determined for each data packet individually. The routers at the border of an AS use specific protocols to 'advertise' the IP addresses inside their AS they can route traffic for. A transit network also advertises routes for addresses that do not reside inside the AS of the transit network operator, but are reachable *through* the transit provider's network.

A so-called *tier 1* transit network is a network that is able to route traffic to virtually all publicly advertised addresses. While the tier 1 networks provide the convenience of being able to reach everyone on the internet, their services are typically expensive. Internet service and content providers will generally try to deploy direct interconnections between their networks in order to reduce the amount of traffic that has to go through tier 1 transit networks. Such an interconnection can happen bilaterally or at an *internet exchange* with multiple networks at once. An interconnection over which traffic is exchanged between two networks without charging for the traffic is referred to as *settlement-free peering*. A network may combine various forms of interconnection (transit and peering) to achieve global connectivity (possibly with redundancy – i.e. when a peering interconnection fails, the network can fall back to the transit network).

3.1.3 Private networks

A private network is a network that is operated by a single network operator, and does not provide network services to other networks. Depending on the type of organisation operating the network, the network can be either of the following:

- *Internet-facing private network***.** These networks are usually operated by media and internet service companies, who generate and own content and distribute it over the internet. These networks exhibit very large amounts of outbound traffic, but (usually) limited amounts of inbound traffic.

- *Internal private network***.** These networks are usually operated by larger corporates and governments. The network supports the internal infrastructure, and also provides internet connectivity to make services outside the network available to users connected to the network. As the network does not provide services to other networks but rather consumes services from other networks over the internet, the traffic volume is higher inbound than outbound. While internal private networks are usually connected to the internet, connectivity may be limited. Usually, hosts inside an internal private network are not visible to the internet, but the hosts inside the network can reach almost all internet destinations.

Private networks usually interconnect using transit, or directly with other private networks if this makes sense from a business perspective (e.g. because two organisations need to work together), or to save on transit traffic. The latter is usually the case for content-producing networks.

3.1.4 End-user and public ISP networks

As operating a private network and the necessary interconnects is quite expensive, smaller organisations typically rely on the services of *internet service providers* to obtain connectivity to the internet. These smaller networks are built on the same technology as all the other networks on the internet, but are much more lightweight. Instead of multiple interconnections to different networks, they usually have only a single interconnection to the network of an internet service provider, who acts as a transit provider.

The ISP handles all routing and other network services for the end-user network. Typically, residential networks only have a single address that needs to be routable on the internet.[21] Due to this, end-user networks are generally not recognized as a separate AS.

Table 1 provides an overview of the different network types and their relevant characteristics.

[21] Most residential networks use a special type of router that allows multiple hosts on the internal network to connect to hosts on the internet, while externally using only a single routable internet address; the router 'fakes' that it is a single host on the internet from which all traffic originates (network address translation). The router maintains a mapping table in order to deliver the right packets to the right internal host.

Dimension	Transit networks and IXPs	Private networks	Public ISP networks	End-user networks
Environment	Field, data centre	Data centre	Field, data centre	Customer premises
Coverage	Intercontinental	International, national	National	Single location
Number of nodes	Low	Medium	High	Very low - low
Traffic volume	High volume Low variance	High volume High variance, controllable	High High variance	High High variance
Traffic type	Homogenous (layer 2)	Heterogeneous	Heterogeneous	Heterogeneous
Network service demand	Limited (routing info, monitoring)	High, homogenous	High (esp. customer-facing), heterogeneous	High
Examples of owners	Level3, Cogent, NTT	Google, Amazon, Akamai, RTL, large corporations	KPN, Ziggo, Vodafone, T-Mobile	Consumers, SMEs
Primary driver	Quality (bleeding edge)	Quality and price	Price	Price
Culture and perspective	IT/Telecom	IT	Telecom	IT

Table 1 Different types of networks and their defining properties

3.2 Opportunities for SDN and NFV

3.2.1 Transit networks and IXPs

Transit networks and IXPs provide a relatively simple, but high-capacity service to their users, which is the delivery of large aggregate streams of traffic to a small set of locations (which are sometimes geographically far apart). Links on transit and IXP networks are typically defined at the data link layer (layer 2) and are largely static over time. Typically, the capacity available on a link is fixed (either limited artificially or limited by the hardware used). The IXP or transit operator usually provides a service that is 'as fast as possible' given state-of-the-art hardware, and does not perform any adjustments to the traffic regarding quality of service.

As these networks primarily deal with large, aggregate, opaque traffic streams, they do not typically offer any higher-level network functionality. IXPs do however typically operate a so-called *routing server*, which collects and distributes information on the routes provided by the networks connected to the internet exchange. This service improves the efficiency of the exchange of this information, as without it, all networks would have to talk to every other network to collect it.

Customers of transit and IXP networks typically do not require detailed control over the service provided. Typically, a network connects to an IXP or transit provider by means of a physical interconnection realised in a data centre where both networks are present. The only administrative matters that need to be sorted out between the transit network and the end-user network is enabling or disabling specific ports and setting up the correct MAC accesses and VLAN tags (for Ethernet-based networks). Some IXPs (such as AMS-IX) already offer programmatic ways for customers to make these changes.

Resellers of traffic on transit networks and IXP typically do provide higher-level network services, and will perform QoS between customers. For these organisations, SDN and NFV may be of relevance in order to provide more fine-grained and flexible QoS functionality. There are also IXPs that operate at higher network layers (e.g. so-called GRXes that exchange mobile data traffic) which may benefit from SDN and NFV to a greater extent – as they have smaller scale, they may be virtual customers on larger infrastructures, which may be more cost-effective than running their own.

3.2.2 Private networks

Private networks are especially suited for application of SDN and NFV technologies and practices due to their specific characteristics.

First of all, private networks are, by nature, typically controlled by a single entity. Therefore, it is easier to deploy network-wide changes to these networks than it is in other networks. As these networks typically serve a single purpose and single set of users, the holistic control aspects of SDN and the efficiency and flexibility aspects of NFV are of great interest.

For private networks, an integration between SDN/NFV and other forms of virtualisation (e.g. computing) makes a lot of sense. Typically, operators of large computing infrastructures have already moved to virtualized computing infrastructures, and would greatly benefit from integrating also the network aspect into their administration and management systems.

3.2.3 Public ISP networks

A key characteristic of public, ISP networks is that they have to deal with large amounts of traffic, with a high variance in the volume. They typically have different types of customers, each with their own requirements regarding the quality of service – being able to quickly provision while managing this complexity is a major challenge for ISPs. These characteristics reflect on the architecture of the networks of ISPs. ISPs typically have a *core* network that is designed to accommodate traffic that is high in both volume as well as variety. The *edge* of an ISP network is designed with completely different architectural goals, the primary one being to provide network access to a highly heterogeneous group of customers in the most cost-effective way possible.

Core network

The core network of an ISP needs to be able to accommodate high volumes of different traffic types and meet the requirements set out in SLAs with customers, even under peak load conditions. For this reason, ISPs typically have specialised, expensive equipment in their core network, because of the features required to meet this type of demand. The specialised equipment often contains features that the ISP isn't using, but still pays for. Nevertheless, the core equipment is not a prime candidate for migration to NFV, as the performance requirements are very high.

Edge network

Given the above, we expect it is more likely for SDN and NFV to start being adopted near the edges of ISP networks. The specialised hardware used in the core is simply more difficult to virtualise than the more generic hardware near the edges of the network. Due to the number of devices, the advantages of centrally managing devices is also greater near the edges of the network than it is in the core. A disadvantage is that the equipment is not centrally located in data centres, which makes a migration towards SDN or NFV more difficult as hardware will have to be replaced at some point in time.

A second opportunity for ISPs regarding SDN and NFV is at the very edge of their network, where they have equipment at the customer's premises. ISPs have ownership over CPEs (modems) and can use this position to provide new services. See for example the Ziggo WifiSpots initiative, which is effectively a completely virtual wireless network with nationwide coverage [42]. NFV and SDN can make deploying such services significantly easier.

NFV and SDN are also relevant for *edge computing*, where services are provided from locations close to end users (e.g. near the DSLAM, or mobile base station).

3.2.4 End-user networks

End-user networks are usually highly heterogeneous and relatively small, especially for consumers. The number of nodes as well as the low complexity of the current set-up does not seem to lead to opportunities for SDN for the end-users themselves.

Any application of SDN is likely only the result of ISPs starting to apply these principles in their access networks. For example, if ISPs start to deliver multiple virtual networks towards customers, then SDN may be needed to direct these traffic streams to the right devices inside the customer network as well. Imagine for instance the creation of a virtual network between a set-top-box and an IPTV service. [22] This scenario may become relevant with the introduction of more Internet of Things applications.

Second, NFV makes sense for easy provisioning of managed services (e.g. firewalls) provided by service providers to customers. ISPs typically already include security solutions (e.g. firewalls, anti-virus and malware scanners) in their product portfolio, but they are either fixed in specific locations in the network, or not integrated at all (e.g. offered as separate software-based solution for the customer to install). Using NFV, ISPs may offer tailored solutions which can be offered 'as a service' and on demand to customers.

3.3 New propositions

SDN and NFV provide technical opportunities for creating new propositions. In this paragraph, we list the propositions that are currently recognized in the market, and are generally seen as interesting and feasible. The list is obviously not complete with respect to future propositions. In addition, admissibility under the (current) European net neutrality legislation was not considered, and the terminology used in the list is (purposefully) not aligned to (current or future) regulatory frameworks.

3.3.1 Specialised access

Internet service providers can provide specialised network access for services that have specific requirements regarding QoS. Several examples:

- Utility companies could connect smart meters to existing internet connections. This could be substantially less costly than the current solution, which is usually wireless (in a reserved frequency band using CDMA, or using a public mobile network).

- Managed appliances for offices (e.g. printer multifunctionals, security systems) and homes (cars) could be connected and controlled.

- Healthcare applications, such as monitoring equipment, alarm buttons, and telepresence systems.

- Payment terminals.

As noted before, many of these applications could also work fine using over-the-top connectivity. The reliability and availability of a software-defined virtual network can of course never be better than can be provided by the underlying link. Virtualisation does provide a clean separation between these (typically sensitive) applications and internet traffic, and allows provisioning a certain guaranteed quality-of-service.

An example of a specialised, virtual access network that is currently deployed is the Ziggo Wi-Fi spots network. In 2013, Ziggo pushed a firmware update to the cable modems at its customer's premises, which made the modem broadcast a second, virtual wireless network named 'Ziggo'. Ziggo customers can connect to this network everywhere it is available – the Wi-Fi network is 'virtually everywhere'. Traffic generated by the guest network is limited and transmitted using a separate virtual network connection (and does not impact the bandwidth available to the subscriber). [42]

[22] Currently, this is typically done by assigning a physical port on the modem for the set-top box. This could be made dynamic (e.g. depending on the MAC address of the TV, assign the right VLAN). VLANs are an example of how SDN can set up specific paths for specific traffic flows. However, SDN enables more fine-grained flow (traffic) management if needed.

3.3.2 Separated service access

In the Netherlands, all major ISPs provide separated services – in the case of KPN, telephony and television are in fact IP-based and delivered over a separate VLAN to customers over a DSL or fibre connection. Interestingly, watching TV over a Telfort DSL connection can impact internet service access, because the television stream receives guaranteed bandwidth which would otherwise be available for internet access. [40] On Ziggo's HFC network, broadcast television streams are transmitted using DVB-C, which is not IP-based. Nevertheless, certain 'interactive TV' and 'on demand' features are delivered over IP.

ISPs can use SDN and NFV technology to quickly and flexibly provision additional services to their end users over their existing access network, separate from the internet. This effectively allows them to provide additional services using favourable network conditions.

Box 1 Networks-as-a-Service

Taking this idea one step further, an ISP could (using SDN and NFV mechanisms) offer a Network-as-a-Service (NaaS)-type of service to (business) customers. As such, a company offering a service with 'special' network requirements (i.e. those mentioned in paragraph 3.3.1) could go to an virtual network provider (VNP, which could be an entity part of an ISP) and demand a virtual network.

Two main service models could emerge:

1. *Provider-managed*: the (business) customer can choose between a range of high level options such as network throughput, network attached storage and computing resources, level of security, network latency, et cetera. The VNP provides the user with a virtual network and ensures that the service level specifications are met (e.g. routing decisions are still made by the VNP).

2. *Customer managed*: The customer demands a virtual network with a set of low-level network characteristics defined (e.g. link distance between two locations, node computing power, and node storage capacity). The customer is responsible for operating the virtual network (e.g. making routing decisions).

This could be an interesting new revenue stream for ISPs (or for a dedicated virtual network provider). The first model would be of interest to companies that have specific network requirements but lack expert knowledge on how to operate a network. The second model would clearly be of interest to companies that already have expert knowledge available but lack access to the infrastructure (e.g. OTT content providers such as Google, Netflix, Facebook, etc.). As the second model allows a customer to define its own network policy, this may have an impact on the performance of the service. For example, two companies with exactly the same virtual network may see a difference in service performance because the algorithm used by the first company is superior to that of the second. As such, when monitoring network performance one may come to the conclusion that the first company's traffic is provided with a higher priority.

This raises a couple of questions from a regulatory perspective:

1. What is the regulatory stance towards NaaS type of services?
2. How would regulation impact provider- and customer managed NaaS?
3. How can conformity with regulations be monitored and enforced?

European legislation explicitly requires net neutrality to be observed by operators of public internet access networks. [35] Because of the way this requirement is worded, it applies only to services that are considered *internet access*. This means an operator is allowed to provide e.g. a telephony or television service on their network separately from the internet service, as long as it is properly defined as such. [23]

3.3.3 Flexible value-added internet access services

In addition to creating *secondary* virtual networks alongside the main internet access network, ISPs can also add additional services to the internet access connection using NFV. ISPs currently provide a variety of higher-layer services to their customers, such as those required for network configuration and internet access (DHCP, DNS), e-mail and sometimes web storage and news services.

Using NFV, ISPs could provide additional internet access services to end users, which can be enabled or disabled flexibly. While some ISPs currently already offer filtering and security solutions, these are currently 'one size fits all' – virtualisation allows ISPs to tailor the service provided to specific customers.

For example, an ISP could offer its customers a 'virtual firewall', which inspects and filters traffic between the internet and the customer. The virtual firewall can be configured to the user's specific needs. Another example are parental control services, which restrict internet access for specific users (e.g. children) or blocks certain content (e.g. pornography and violence).

Note that the value added services can also be provided by a service provider outside the ISPs network if the ISP can flexibly route the traffic to and from that service provider (i.e. using SDN). ISPs can also employ NFV to provide these services on virtual CPEs, so that the service effectively executes inside the user's network. A *virtual CPE* is a modem located at the user's premises, on top of which virtual network services can be run, and where the core functionality of the modem is typically also just one of the virtual services running on it. [18] Finally, end users still have the option to implement value added services by installing software on their own computers or adding specific equipment to their networks.

3.3.4 Edge computing services

In the past few years, content providers have optimized their networks to reduce the distance between their network and their customers as much as possible. The primary reason for doing this is to reduce the amount of (expensive) traffic required to serve customers that are located far away from the service provider. The second reason is that shorter distances reduce the delay experienced by traffic between the service provider and customers, leading to a better user experience. Typically, content providers build or lease a content delivery network, which caches the most popular content on servers located close to the end user (typically in the network of an ISP, or near internet exchange points such as AMS-IX).

Using SDN and NFV, ISPs will be able to more flexibly accommodate such 'guests' in their network. In addition, they may be able to open up their network even further, so that content

[23] See consideration (16) and article 3 sub 5 in [35]. The latter states that "*[ISPs] shall be free to offer services other than internet access services which are optimised for specific content, applications or services, where the optimisation is necessary in order to meet requirements […] for a specific level of quality. Providers […] may offer or facilitate such services only if the network capacity is sufficient to provide them in addition to any internet access services provided. Such services shall not be usable or offered as a replacement for internet access services […]*".

delivery nodes can be placed even closer to end users. ISPs could also offer virtualised hosting of network services on the subscriber's modem (CPE). [18]

This type of 'edge' or 'fog' computing can potentially be very helpful for services that have very strict latency requirements, such as online gaming and virtual reality. It can also be used for security purposes, e.g. to prevent DDoS and other digital attacks by blocking malicious traffic at a very early point in the network. Securing the virtual CPEs themselves however may prove to be a difficult issue to overcome. [1]

3.3.5 Fine-grained wholesale access

SDN and NFV technologies enable network infrastructure owners to provide access to their networks at a very low level and in a way that is highly tailored. This ability allows these providers to provide fine-grained access to alternative network operators. The added flexibility can be realised in the following ways:

- As discussed earlier, SDN can be used to flexibly create VLANs to end users. This would allow the guest operator to provide services such as TV and telephony that it would otherwise have to buy wholesale as well from the host network operator. An alternative operator can also more flexibly switch between these services.

- Using SDN, a guest operator can more flexibly migrate from wholesale backhaul and core connectivity to alternative forms of connectivity (i.e. their own or leased fibre infrastructure).

- Using SDN, a guest operator can start to offer lower-level services to its customers, such as layer 2 VLANs/VPNs.

- If the host operator provides the guest operator access to its own virtual CPEs, a guest operator could be able to provide one of the triple play services (television, telephony or internet access) while other services are provided by another operator.

3.3.6 Network integration

SDN and NFV could make it easier for network operators to integrate access networks. For instance, a mobile operator could use SDN and NFV features provided by a fixed network to realise Wi-Fi offloading of traffic from smartphones.

3.3.7 Cloud-based middleware

Cloud providers such as Amazon and Google currently offer a large spectrum of virtualised computing and networking services. These services are used by small to medium-sized companies that have a need for computing and networking and also require their infrastructure to be able to scale up quickly. The networking services provided by the cloud providers however typically stay inside the datacentres of these operators.

Using SDN and NFV, cloud operators such as Amazon and Google could begin to offer networking services to consumers and SMEs that they currently cannot offer, such as load balancers, firewalls, filters and even traffic inspection services.

3.3.8 Temporary specialised access services

SDN and NFV provides network infrastructure providers to quickly provision links between locations. This ability can be used to quickly provide specialised access services.

One application of this is to provide quick failover in the case of emergencies. Consider for instance the connectivity required by an operations centre of a nationally operating company (e.g. utilities). Should there be an emergency at the operations centre, the operator providing connectivity to that location could simply migrate the full set of network connectivity (including any additional services and security configuration) to a different location. Failover services are currently already provided on virtual computation platforms, where services can be moved to different hardware instantly after failure is detected.

3.3.9 Overview

Table 2 gives an overview of the different new propositions discussed above.

From this table, a few conclusions can be drawn. First of all, access network operators appear to have the most opportunities to use SDN and NFV to provide new services. The cooperation of the access network operators is also required for other new propositions to be possible at all. Third parties should be able to make use of the SDN and NFV platforms available on the access infrastructure through an interface similar to what is used by the access network operator itself.

Table 2 Overview of new propositions made possible by SDN and NFV technology

Service	Producer	Consumer	Required interface for producer
Specialised access	Access network operator	Utilities, financial services, industry (larger corporates)	SDN API on access network
Separated service access	Access network operator	Content providers	SDN API on access network, NFV in backhaul network
Flexible value-added internet access services	ISP	Consumers, SMEs	SDN API on access network, NFV in backhaul network
Edge computing services	Access network operator	Content providers	SDN API on access network, NFV in backhaul network and CPE
Fine-grained wholesale access	Access network operator	ISP (using wholesale access)	SDN API on core and access network
Network integration	Access network operator	Other access network operator	SDN API on core and access network, possibly NFV
Cloud-based middleware	Cloud platform provider	Consumers, SMEs	SDN API on access network, NFV in backhaul network and possibly CPE
Temporary specialised access	Access network operator	Utilities, financial services, industry (larger corporates)	SDN API on core and access network, possibly NFV

3.4 Implementation drivers

The implementation of SDN and NFV is driven by two forces in the market. The first is the *dominant paradigm* for SDN and NFV, which is an attribute of the supply side. The second is the *adoption rate* for SDN and NFV technologies, which is an attribute of the demand side. Both drivers will be discussed in further detail below.

3.4.1 Supply side: dominant paradigm

From the side of the network equipment suppliers, there is a strong technologically driven force that leads the vendors to deliver and push their solutions based on SDN and NFV. There is however still a large amount of uncertainty in the market regarding what exactly is considered SDN and NFV, which SDN and NFV based solutions are feasible propositions on the market, and what will be the dominant technological choices for SDN and NFV.

We strongly suspect that SDN and NFV will at some point converge towards accepted standards and definitions. It is however too early to tell whether these will (a) be open standards[24], and (b) what the scope of this standardisation will be.

A second aspect of the dominant paradigm for SDN and NFV is the willingness of network infrastructure owners to provide access to SDN and NFV functionality through an *interface* to third parties. Without such an interface, many of the new propositions made possible by SDN and NFV are impossible to realise, or can only be realised by the network operator itself. The operator may have strong incentives to keep access to the interface highly restricted.

To summarize, the key parameters for this driver are the following:

- **Level of standardisation:** The extent to which the industry will succeed in the creation of a small number of standards, which allow interoperability of network equipment and orchestration layers for SDN and NFV.

- **Degree of standards openness:** The extent to which the standards developed in the industry for SDN and NFV are open. This determines the extent to which new suppliers will be able to introduce equipment to the market based on open standards.

- **Degree of network openness.** The extent to which network infrastructure operators are willing to provide interfaces to SDN and NFV functionality to third parties.

Note that virtually all of these parameters are determined globally, except when specific regulation is made to steer either in a particular direction for the Netherlands.

Box 2 Open initiatives for standardisation of SDN and NFV

In the context of openness and SDN, OpenFlow cannot be ignored. OpenFlow is a protocol for communication between a network controller and a network switch or router. [27] It allows the network controller to configure the paths that different types of data packets ('flows') follow over the data plane of the switch or router. As a standardized protocol, OpenFlow abstracts away differences between different types of hardware from different vendors.

The OpenFlow standard is managed by the Open Networking Foundation (ONF). The foundation is funded by several market parties, including Facebook, Google, Microsoft and Verizon. A large number of vendors has indicated that they have implemented or will implement OpenFlow support in (a subset of their) products, including Cisco, Huawei, Juniper, Brocade and Arista.

Another open source project is OpenDaylight, which is an open source controller platform for SDN, and can make use of the OpenFlow protocol to control individual switches and routers. [31] The OpenDaylight project is run by The Linux Foundation.

In addition to the open platforms for SDN and NFV, there are several different closed implementations of similar functionality. Note that even while different open standards exist, vendors can choose to implement only a subset of open standards, and provide integrated, closed solutions for different parts of the technology stack.

[24] Most likely, the standards will be made publicly available without charge (but under copyright). These standards define open interfaces (such as OpenFlow), the implementation of the standard in actual products (e.g. software) may however happen in a vendor specific way.

3.4.2 Demand side: adoption rate

On the side of network operators, the main driver is a combination of cost/efficiency and flexibility, which leads the operators to implement or migrate towards SDN and NFV technologies in their networks.

From a technical point of view, SDN and NFV can make networks more efficient and flexible. This has an evolutionary rather than a revolutionary character: many of the things that SDN and NFV make more efficient or flexible are already possible with current networking technology. For SDN and NFV to have an impact, they hence need to provide either a significant technical advantage, or a substantial improvement in organisational efficiency.

Technical efficiency improvement

From a technical point of view, the adoption rate will be driven by the extent to which SDN and NFV technologies and offerings in the market can realise improvements in technical efficiency, resulting from higher flexibility in network management and configuration, or in overall better-performing networks.

The main question to be answered here is whether SDN and NFV will be able to provide a significant added value atop of current technological possibilities. As we have seen earlier, a lot of the applications touted as being the ‘killer features’ of SDN and NFV are actually already possible using existing equipment and software (albeit more difficult to implement from an organisational point of view). The question to be asked for each of these features is therefore how much easier SDN and NFV make deploying the functionality.

Three ways in which SDN and NFV may be able to provide the biggest benefits are the following:

- The top-down view provided by SDN and NFV allows holistic management of traffic streams, enabling network capacity to be utilized more efficiently. However, in many networks capacity is not an issue.

- Top-down configuration allows for more flexibility and agility. Operators may be able to use links they previously couldn’t use for a particular application (e.g. due to quicker provisioning, the ability to temporarily route other traffic over other links, et cetera).

- SDN and NFV may provide new ways for monetizing a network. See paragraph 3.3 as well as Box 1 for a discussion of possible new propositions.

Organisational efficiency improvement

If SDN and NFV cannot provide new functionality, they need to provide a significant organisational benefit in order to be adopted at all. Such organisational efficiency improvement can be achieved in the following ways:

- **Lower the effort needed to provision networks and links**. Using SDN, network equipment can be configured centrally and from a single interface, reducing the number of tasks that are required for a specific network device.

- **Streamlining of procedures**. Individual configuration of switches and routers is reduced to a minimum. Current users of SDN however indicate they still often need to go to the individual boxes to troubleshoot issues. It is unclear if SDN will be able to improve so that even this type of individual access is no longer required.

- **Improved resiliency, as failures can be dealt with more quickly.** Failure in the SDN orchestration layer may however lead to inability to configure the whole network.

- **Top-down definition and verification of network policies** (e.g. regarding security, Chinese walls, etc.). Note that while SDN reduces the attack surface if direct access to node configuration is closed, this is unlikely to happen in practice, as low level access to network devices is still desirable for troubleshooting purposes. The SDN orchestrator itself is a new, large and interesting attack surface – control over the orchestrator implies control over the full network, which makes the potential impact of an attack much larger.

Some interviewees indicated to us that while SDN and NFV do provide significant benefits from a network administration point of view, they also observed a shift of workload from network administrators to software engineers, who are tasked with designing the management systems that control the SDN and NFV-based platforms.

The key parameters of this driver can be summarized as follows:

- **Efficiency**: to what extent will general purpose hardware be able to compete with special purpose hardware, and for which applications?

- **Flexibility**: how much and for which applications can SDN and NFV improve flexibility to a point that justifies investment in SDN/NFV?

- **Migration path**: Is there are smooth migration path for existing (traditional) networks to move to SDN/NFV?

3.5 Implementation scenarios

The different scenarios for SDN and NFV deployment follow from the two drivers discussed in the previous section: *dominant paradigm* (supply side) and *adoption rate* (demand side). Figure 7 shows a matrix with the two drivers presented on the horizontal and vertical axis. Each quadrant of the matrix represents a distinct scenario for SDN and NFV deployment .We will discuss the different scenarios in more detail below.

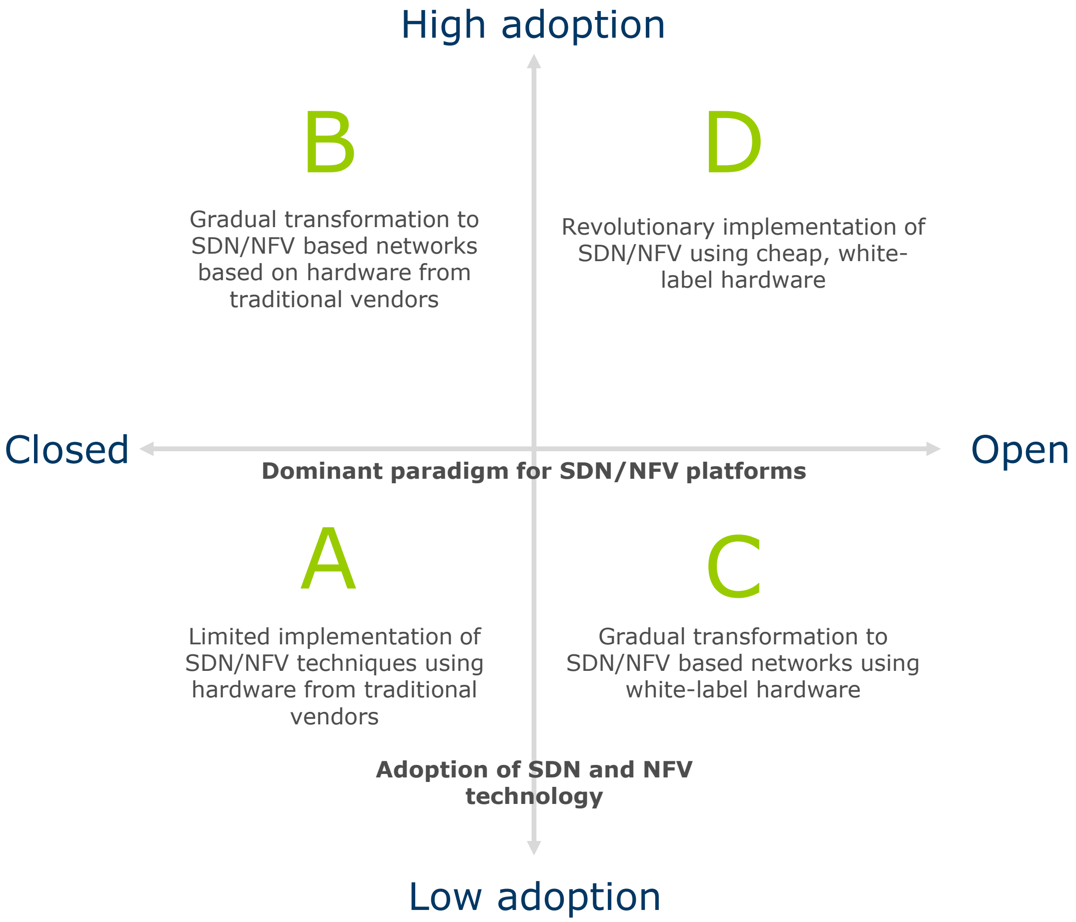

High adoption
B
Gradual transformation to SDN/NFV based networks based on hardware from traditional vendors
D
Revolutionary implementation of SDN/NFV using cheap, white-label hardware
Closed
Open
Dominant paradigm for SDN/NFV platforms
A
Limited implementation of SDN/NFV techniques using hardware from traditional vendors
C
Gradual transformation to SDN/NFV based networks using white-label hardware
Adoption of SDN and NFV technology
Low adoption

Figure 7 Overview of scenarios for SDN and NFV deployment

3.5.1 A: Limited implementation of SDN/NFV, traditional vendors

The first scenario is closest to the current state of the market. In this scenario, SDN and NFV technologies will be adopted to a limited extent, and will be added to current network equipment offerings rather than be offered as completely new products and services. The network equipment vendors in the market more or less keep their current position, and will compete with each other to provide value added services by adopting SDN and NFV principles.

3.5.2 B: Evolution towards SDN/NFV, traditional vendors

In this scenario, SDN and NFV will be adopted more widely than in the first scenario. However, the transition is still rather slow and happens smoothly, as SDN and NFV technologies are added to existing offerings and become part of infrastructure upgrades of the network operators. Usage of SDN and NFV will largely be focused inwards.

3.5.3 C: Evolution towards SDN/NFV, new vendors

In this scenario, new players will enter the market with new, SDN/NFV based solutions. Networks will slowly adopt the new hardware as part of modernisation, replacing traditionally ‘large’ network equipment with the smaller, more flexible and typically cheaper networking equipment. The transformation will start in networks that have extremely high requirements, and will then trickle down to the more conservative networks.

3.5.4 D: Revolution of SDN/NFV, new vendors

In this scenario, the introduction of cheap, flexible hardware solutions for SDN and NFV leads to a shock in the market, where networks are rapidly adopting the new technology to be able to reap the benefits of both the technology as well as a significant cost reduction for equipment.

Note that these scenarios reflect possible "outcomes of the world". It is beyond the scope of this type of scenario analysis to explore the factors that influence which scenario will materialize. Hence, for the purpose of this study, the scenarios represent exogenously determined states of the world. The aim of this approach is to make policy makers aware of fundamental uncertainties regarding the direction and development of network virtualisation. Nevertheless, we tried to articulate the key parameters underlying the drivers representing the two axes (for the supply side, and for the demand side). Moreover, because of the fundamental uncertainty that is involved, it is not possible to identify what policy makers or governments can do to influence the coming about of a specific scenario.

4 Economic impact

This chapter explores the economic impact of network virtualisation. The economic perspective is partly based on the framework proposed by Bennett et al. (2001), which serves to analyse, in a consistent and systematic way, 'future' policy issues related to market developments that are to some extent exogenous to policy makers.[25] We adapt the framework to account for the fact that the technological developments regarding NFV and SDN are, in the current research set-up, a dominant force, and in that sense, a starting point for the exploration. To assess the economic impact of this technological change, we will assess its impact on static and dynamic efficiency. These welfare indicators depend on 'efficiency drivers' such as market structure, the nature of competition and public policy. The scenarios that were developed in the previous chapter will help to deal with the uncertainty that is involved in assessing the impact on welfare.

Section 4.1 recapitulates and explains the economic concepts. Section 4.2 discusses them in the context of network virtualisation, building on the previous chapters. Section 4.3 explores the impact of network virtualisation on static and dynamic efficiency, and briefly discusses the effect on the economy and society as a whole through spill over effects. In doing so, we try to distinguish the possible impacts among the scenarios.

4.1 Economic concepts

From an economic perspective, policy may be designed such that welfare, that is, the sum of consumer’s surplus and producers surplus, is maximized, while safeguarding any 'public interests' that are valued by society. Consumers surplus measures the aggregate net benefits (i.e., the utility level from using a good or service, minus the price paid for it) of consumers participating in a market. Producers surplus measures the aggregate variable profits (where variable profits equal revenues minus variable costs) of the firms in a market.

The underlying idea is that it makes sense to maximize the total size of the 'pie' that can be divided among market agents. If competition is effective, consumers will automatically get a 'fair share'. If, furthermore, there are no substantial market failures that undermine specific 'public interests', a market can be said to be functioning well.

In Box 3 below, we explain how the welfare notion can be made operational.

[25] [15] presents the economic framework of [7] in a more concise way.

Box 3 Operationalisation of the notion of welfare

The maximization of welfare corresponds to the notion of efficiency. Consider an 'allocation', which may corresponds to the way inputs (e.g. materials, labour, capital) are used to produce output (goods and services), or to the matching of supply and demand in a market. An allocation is said to be ('Pareto') efficient if there exists no other feasible allocation that is preferred by unanimously by all participants. For policy purposes though, one may use a more practical definition: efficiency is obtained by an optimal matching of supply and demand, that is, in accordance with consumer's preferences, such that goods and services are produced at the lowest possible cost. Welfare is maximized if a market works efficiently, or is 'well-functioning'.

There is an important distinction between the short and the long run:

- In the short run, technologies are assumed to be given, or fixed. Hence short-run efficiency refers to a situation in which there is an optimal matching of supply and demand while existing goods and services are produced at the lowest possible cost, making use of existing technologies.

- In the long run, new technologies may emerge, resulting in new goods and services, or in cheaper ways of production. Hence, long-run efficiency refers to a situation in which there is an optimal matching of supply and demand resulting from the introduction of new goods and services, more variety, as well as production technologies that reduce cost levels.

Based on this distinction, instead of short-run and long-run efficiency, one can speak of static and dynamic efficiency:

- Static efficiency refers to welfare in the short run. It is obtained by an optimal combination of given inputs, subject to the constraints imposed by existing technology. It is maximized by making the best use of existing resources and technologies, and hence by abstracting from innovations and investments. Higher static efficiency arises from an improved allocation of inputs and an improved matching of supply and demand.

- Dynamic efficiency refers to welfare in the long run. It is obtained by an optimal combination of inputs over time that maximizes the present value of current and future welfare, by allowing for investments and new technologies. It is maximized by allowing for innovations and investments. Higher dynamic efficiency arises from process innovations, leading to lower costs of production, as well as product innovations, leading to new or improved products or services that consumers value more than existing ones.

Static efficiency can be seen as a situation in which consumers get good value for money while (and because) competition is effective. It is a combination of allocative and productive efficiency, taking investment and innovation levels as given. The relationship between competition and static efficiency is typically straightforward: in the short run, more competition is typically 'good', as it encourages firms to cut slack, and creates downward pressure on prices, which reduces the 'deadweight' welfare loss.

Dynamic efficiency refers to a market's capacity to create and adopt new technologies, both at the process and the product level. Such developments take time to come to fruition, so consumers may not immediately benefit from increased variety at low prices. Hence the relationship between competition and dynamic efficiency may not be obvious. For low intensities of competition, firms have little incentives to invest in innovation. If competition becomes more intense, they start feeling pressure to leapfrog ahead of their competitors, by innovating. This is the 'escape-competition' effect. For higher levels of competition, however, this effect may be reversed into a 'Schumpeterian' effect, meaning that more competition reduces firms' incentives to innovate. When competition is very intense, firms become too pessimistic about recovering investments, which reduces their incentives to invest. Arguably, there is an 'inverse-U' shaped relationship between competition and investment, which implies an optimum level of competition: too little, as well as too much competition is sub-optimal. [2] Hence, according to economic theory, the incentives to invest and innovate depend on having sufficient possibilities to earn monopoly rents. These rents should not be too high, however, to avoid the common distortions of monopolies.

In practice, it is difficult to measure or estimate welfare, due to a lack of empirical data. For instance, to measure consumer's surplus, one must know consumer preferences, demand parameters and demand elasticities. And to measure producer's surplus, one must have information about profit levels and investments. For practical purposes, when such parameters are unknown, one may use (qualitative) proxies for static and dynamic efficiency. To assess static efficiency, one may use proxies such as the intensity of competition, the number of competing firms, transparency for consumers, and the pressure to eliminate slack in production processes. For dynamic efficiency, relevant proxies are investment levels in R&D, the number of product innovations, the number of process innovations, the number of patent applications, increases in variety for consumers, increases in the quality of goods and services, and the possibilities for market entry.

To perform a quick scan of static and dynamic efficiency, one may distinguish the following 'efficiency drivers': (1) market structure; (2) anti-competitive practices; (3) public policy; and (4) technology. [7] In this research, the latter efficiency driver will receive a special treatment, due to its fundamental, overarching impact on the other drivers.

4.1.1 Efficiency driver 1: Market structure

Market structure and the way it may be changed by entry and exit determine, to a large extent, the nature and intensity of competition. Hence they strongly affect static and dynamic efficiency. The most important characteristics of market structure are: the number of suppliers and their market shares; concentration tendencies (such as M&A activities); the presence or absence of entry barriers; vertical integration versus separation; and the way in which firms compete. The main dimensions are (1) horizontal, referring to firms at the same level of the value chain; and (2) vertical, referring to firms at different levels of the value chain. The first dimension pertains to competing firms, and the second one to firms in a supplier-user relationship.

4.1.2 Efficiency driver 2: Anti-competitive practices

Anti-competitive practices, such as cartel behaviour and foreclosure, may reduce the intensity of competition, and harm both static and dynamic efficiency. A reduction in competition may, under certain circumstances, lead to more innovation. Nevertheless, in general it is safe to assume that anti-competitive practices are harmful for welfare, both in the short and in the long run.

4.1.3 Efficiency driver 3: Public policy

Public policy, in particular in the form of competition policy and sector-specific regulation, may have a large impact on the effectiveness of competition, as well as on the safeguarding of public interests. Hence, the effectiveness of competition policy and regulation is crucial for static as well as dynamic efficiency.

4.1.4 Efficiency driver 4: Technology

This efficiency driver ultimately determines the value that goods and services deliver, the underlying cost levels, and the range of potential applications. Note that even when the resources for innovation are in place, coordination problems and lock-in effects may prevent superior technologies from being adopted in the market. Recall that this efficiency driver will be given a special treatment, due to its overarching impact on the other efficiency drivers (see the next subsection).

The list of efficiency drivers may not be complete, but generally provides an adequate picture of the possible factors that affect static and dynamic efficiency. They can be used to assess the most relevant characteristics of a market (e.g. current players, newcomers, nature of competition), the institutional environment (legislation and regulation) and the technological constraints that affect static and dynamic efficiency.

4.2 Technological change as a primary force

As was remarked above, the fourth efficiency driver gets priority attention, due to the overarching impact on the other drivers. With respect to technology, the following types of developments may be particularly powerful to drive technological changes in networking sectors, while impacting the first three efficiency drivers: [12]

- separation of data and control planes;
- open control interfaces for network devices of different vendors; and
- programmable control.

The virtualisation of capabilities (network resources, network protocols, computational resources and storage resources), combined with SDN, allows for the delegation of management and configuration tasks to third parties, which in turn allows for more differentiation and the development of new business models. [25] They will have an impact on market structure and the nature of competition, and possibly also on public policy. For example in 5G mobile networks, an increasing demand for traffic, heterogeneity in wireless environments, and diverse service requirements increase the complexity of network management. SDN then allows for a very large number of devices and connections, as well as high bandwidth and low latency. [12] Similar developments take place in fixed networks, where SDN makes it possible to connect, via public networks, two or more locations with specific characteristics that support business critical and mission critical applications.

More generally, as described by the notion of the "Fluid Internet" [25], a virtualized future internet will provide the ability to dynamically manage services in an end-to-end way (see Box 4). Software-oriented design in telecommunications networks is different from SDN for the internet, as the latter network mainly deals with packet forwarding. Nevertheless, the main concepts from SDN for the internet (decoupling of data and control planes, the use of logical centralized control to manage forwarding in large networks) may serve as a reference model for network design. [12] We will come back to these developments in section 4.3.

Box 4. The Fluid Internet

While its original design implied a best-effort routing system, the Internet has gradually evolved into a more service-centric delivery platform. Services have become more interactive, applying rich functionality (e.g. based on awareness of context) and making sophisticated use of underlying content. Due to these developments, delivery requirements have become more stringent. Nevertheless, management of the Internet has remained relatively static, at the cost of the development of dynamic end-to-end guarantees for service delivery.

Latré et al. propose the notion of the Fluid Internet to address these management challenges: *"The Fluid Internet seamlessly provisions virtualized infrastructure capabilities, adapting the delivery substrate to the dynamic requirements of services and users, much like a fluid adapting to fit its surroundings. As such, the Fluid Internet gives a service provider the ability to manage its services end-to- end and elastically."* ([25] p. 1)

This paradigm is based on a combination of notions related to network virtualisation, cloud computing, and service-centric networking. It addresses end-to-end quality guarantees in a dynamic environment of users with changing and varying service requirements. Basically, the Fluid Internet allows for dynamic management of services in an end-to-end manner.

In the developments explored in this study, the fourth efficiency driver, technology, is a dominant force, with an overarching impact on the other three efficiency drivers. Therefore we treat it separately: chapter 2 described and explained the technological aspects of network virtualisation. Given that a technological development is the primary focus, these aspects cannot be seen separately from market structure, anti-competitive practices and public policy. In the following sub-section, we will discuss the impact of NFV/SDN on the remaining efficiency drivers.

4.3 Developments and economic impact

To discuss the impact of NFV/SDN on the efficiency drivers market structure, anti-competitive practices, and public policy, we will relate these efficiency drivers to the implementation scenarios presented in chapter 3. Based on that, we will be able to address potential implications for static and dynamic efficiency.

With respect to vertical relationship between firms and customers, it is useful to provide a sketch of the value chain, even though in a market like this, there does not exist a single, fixed set of layers among market participants. Also, to add some focus, we simplify the value chain by abstracting from transit networks, internet exchange points, and private networks (see chapter 3). Hence, most of the discussion that follows applies to the pivotal level in the value chain of end-user and public ISP networks. Therefore, at a stylized level, one can distinguish:

1. Hardware suppliers ('vendors'), providing hardware to network providers and OTT providers who make their own investments in infrastructure. Different types of hardware vendors offer dedicated, specific network equipment as well as cheaper, generic (commodity, or 'white-label') equipment.
2. Software developers, developing networking software. Note that it may be the case that hardware developers perform software development in-house.

3. Network providers, operating the infrastructure for electronic communications (telecommunications, internet access, media content, OTT services).

4. Alternative and virtual (telecommunications services) providers, who may invest in their own (partial) infrastructure, and need to purchase wholesale access to (specific parts of) infrastructure, such as local access networks, from incumbent operators. Hence, some of these providers invest, to a certain extent, in their own infrastructure, while others resell existing services under a different brand name.

5. OTT providers, using network operators' infrastructure to provide content and services to end-users (typically subscribers to network providers, and alternative or virtual providers).

6. End-users of networks and services and content that run over these networks. One may distinguish residential consumers (including small and medium enterprises, typically on a single physical location) and (larger, typically with multiple premises) corporate customers.

Scenario A (low adoption, closed paradigm) can be seen as a baseline scenario, since it depicts, to a large extent, the current situation, in which so far, there has been little adoption of SDN/NFV techniques, while SDN/NFV platforms are still relatively closed (more on that below).

We briefly recapitulate some observations made in chapter 2, to assess this scenario as well as current market forces and developments.[26] Scenario A is characterized by a limited implementation of SDN/NFV techniques, using hardware from traditional vendors. Many OTT providers do not only offer 'pure' OTT services, but are involved in infrastructure as well. For instance, they may operate their own data centres (DCs) and content delivery networks (CDNs). More generally, DCs currently face inflexibility regarding adapting network equipment, while conventional network equipment is relatively expensive, due to the bundling of hardware and software.

The nature of competition between network operators involves a strong pressure to increase bandwidth for end-users at decreasing prices. Without bundling of access services and content (e.g. through vertical integration with media companies),[27] network operators have little scope for horizontal differentiation.[28] In such a situation, the access services offered by network operators can be seen as commodities, possibly vertically differentiated on the basis of the speed of connections.[29] The pressure on network operators due to commoditization is even stronger in the light of the growing importance of OTT providers (as, for instance, illustrated by the success of Google, Netflix and WhatsApp). They typically offer more innovative content and services than network operators, often while being able to charge higher mark-ups, or by creating leverage through a multi-sided market strategy. On the other hand, one could argue that OTT providers were successful in undermining comfortable market positions of established players, in markets without fierce competition.

[26] See also [14].

[27] Vertical integration refers to M&A activities by parties at different levels in the value chain.

[28]Horizontal differentiation refers to differences in product characteristics that are appreciated differently by consumers with different tastes, and that cannot be ordered in an objective way. Hence it is not possible to make statements saying that one product is 'better' than another one.

[29] Vertical differentiation refers to differences in product characteristics that are appreciated in the same way by different consumers, so that products can be ordered according to objective quality levels. Hence it is possible to make statements saying that one product is 'better' than another one.

Hence network operators have been facing a decrease in revenues, due to the growth of OTT providers. At the same time, traffic growth creates a need to invest in (fixed and mobile) networks, to increase the speed that end-users experience. However, network operators may invest in their networks without being able to recover these investments through higher mark-ups, in particular if they face strong price competition.

Access networks form the most expensive parts of telecommunications infrastructure. They often comprise various types of network equipment, such as modems, local access networks (cable, DSL, 3G/4G mobile, FttH), and the equipment to support connections to end-users' premises (e.g. optical equipment in the case of FttH networks). Providing a mixed infrastructure requires multiple types of access technologies, and is therefore costly. Such networks are complex to operate, as well as inflexible regarding service innovation.

Summarizing, network operators are facing (1) inflexibility regarding the functionalities that they offer; (2) complexity due to legacy technologies and systems; and (3) high cost levels originating from design choices in infrastructure. On top of that, OTT providers have obtained strong market positions as well as strong relationships with end-users, who are able to use OTT services by using ISPs' networks. Moreover, in scenario A, there is still little standardization of SDN/NFV techniques, hardware is relatively closed and dedicated, third parties have little possibilities to create virtual networks for end-users on top of existing networks owned by incumbent network operators, and there is little cooperation between network operators and OTT providers. All these aspects correspond to a 'closed' paradigm.

The situation depicted above, which corresponds to (baseline) scenario A, implies baseline levels for static and dynamic efficiency. We will compare how the scenarios B, C and D may perform with respect to static and dynamic efficiency, relative to scenario A.

In the short term, operators may use various fixes to control the cost of new network deployments, such as data offloading to Wi-Fi from LTE and constraining end-users' download volumes. If legislation would allow it, operators could use DPI to shape traffic (note that this is not allowed though). To reduce transit costs, they may cache content at the edges of networks — which may also save costs for OTT players and improve QoS (e.g. Netflix). In the case of mobile networks, operators make additional revenues by providing access to MVNOs (while wholesale tariffs are not regulated), although typically based on long-term and rigid contractual agreements. Due to the nature of these agreements, MVNOs have little possibilities to shape the services they offer to end-users. For fixed networks, access regulation may limit the profitability of providing access.

In the longer term, without excluding the possibility that the market may remain stuck in scenario A, we will consider the possibility that the market will migrate to another scenario in the coming, say five, years. Network operators are under pressure to come up with more innovative solutions, at least to control CAPEX and OPEX, and also to introduce new functionality and services to end-users, by using network resources in more flexible and efficient ways. The developments of network virtualisation, NFV and SDN (as described in chapter 2), may create opportunities to do exactly that, and while doing so, change the relationships with other players in the value chain, such as suppliers and OTT providers. The degree and way in which this happens, is depicted in scenarios B, C and D, relative to baseline scenario A. Below we will explore how this may affect static and dynamic efficiency. In addition, we will discuss how network virtualization affects the economy and society as a whole, through positive spill over effects.

4.3.1 Efficiency driver 1: Market structure (horizontal and vertical relationships)

Here we discuss the impact of market structure developments in scenarios B, C and D. Many developments are relevant in all scenarios, but in different degrees. We will come back to these relative differences below.

Regarding hardware vendors, some traditional vendors may develop from selling bundles of specific hardware and dedicated software to selling specialised software that runs on general purpose hardware. These developments tend to increase static efficiency.

In addition, the possibilities for software developers to enter the market for telecommunications services increase, because their expertise increasingly becomes a core element of the infrastructure. To do so, they depend on network providers for access to infrastructure. As a consequence, the intensity of competition among alternative and virtual service providers may increase. This will have a positive impact on static efficiency. The impact on dynamic efficiency is uncertain. Due to increased competition in the services market, dynamic efficiency may decrease or increase, corresponding to the inverse-U shaped relationship between competition and innovation. In their identity as software developers, however, these suppliers may get stronger incentives to innovate in the direction of telecommunications architecture and services, as it enables them to enter this market.

OTT providers currently already invest in data centres (DCs) and content delivery networks (CDNs), allowing them to get more control over the delivery of their content and services, and to offer more heterogeneous sets of services to end users. The possibilities to do so will improve due to network virtualisation. As this may invite them to invest more, there may be a positive effect on dynamic efficiency.

DCs face more stringent requirements on their networks to be able to support the increasing heterogeneity of services that run through them. Virtualisation technologies allow them to meet such requirements, by increasing the flexibility of their networks and reducing the network costs. This is likely to increase, at least, static efficiency.

There are various effects for network operators, both at the individual level, and at the level of competitive interaction. First, consider the impact on operators individually — temporarily abstracting from competition.

Effects on operators individually

For network operators, the virtualisation of capabilities (network resources, network protocols, computational resources and storage resources), combined with SDN, allows for the remote management of physical networks and end-to-end service management. Furthermore, it enables flexible scaling of capabilities (network resources, network protocols, computational resources and storage resources) based on service demands. [25] SDN enhances the scope for matching network capabilities by network operators, which translates into a more flexible and faster matching of demand and supply, increasing allocative efficiency through adapting the underlying technology. Whereas the current mode of competition among network operators is often based on stand-alone infrastructures, NFV will allow for more possibilities to 'mix and match' infrastructures and to increase service differentiation. This will have a positive impact on static efficiency, and possibly also on dynamic efficiency, if it leads to more investments and innovation. As was discussed in chapter 3, SDN/NFV can be expected to lead to various new propositions, increasing quality, functionality and variety for end-users. This also contributes to dynamic efficiency.

There are mixed effects on capital expenditures (CAPEX) for network operators. As complex control logic in a network is moved to an external device (to control multiple network devices), capital expenditures increase, and due to additional costs for controllers, line cards, and so on. The possibility to implement traffic management reduces the necessary network capacity, which decreases capital expenditures. Using simpler network devices also leads to cost savings. New open control interfaces and software-defined control will reduce time and cost to reconfigure and optimize access networks, and to introduce new network features. Overall, while it is impossible to tell if capital expenditures will decrease or not, it seems likely that a given level of investment will result in more value for network operators. In other words, static efficiency is likely to increase, while resources may become available for other purposes (alternative investments) which is good for dynamic efficiency.

There are mixed effects on operational expenditures (OPEX) by network operators. The costs of maintenance and repair tend to go down. The cost of service provisioning is reduced as well, due to an automated network configuration. The cost of testing before rollout, and installing network equipment will change: more effort can be devoted to robustness checks, improving the implementation. This shifts costs to an earlier stage, while the total transition cost may decrease. Software-defined control architectures will allow for a more efficient use of networks (infrastructure, and in case of mobile: spectrum).[30] Although the impact on OPEX may go both ways, it seems likely that overall, the operational expenditures go down. Hence, static efficiency is likely to increase.

Whereas current network virtualisation mainly aims at networking capabilities (e.g. bandwidth), future applications can be expected to support flexible allocation of storage and computing resources, as well as more high-level capabilities. In combination with other developments, one may envisage the possibility for end-to-end delivery paths owned by different (physical) network providers, providing varying QoS guarantees that are linked to content as well as service requirements (for instance, SDN can support intelligent and dynamic QoS). The underlying delivery chain can be characterized as a Virtual Service Infrastructure.

Overall, the observations on CAPEX, OPEX, and innovations suggests that network operators will become more flexible and be able to provide more added value, in a more cost-efficient way. Thus, one may expect a positive effect on static and dynamic efficiency.

Consider mobile as an example. SDN-based design of mobile networks may help to tackle difficult problems in cellular and other wireless access networks, to more effectively manage complexity, heterogeneity and consistency in networks, and allowing for further innovations in network architectures. SDN allows spectrum to be managed more efficiently, as the logical centralized control can take spectrum usage into account and implement spectrum sharing. For instance, SDN allows for programmable control to (more effectively) coordinate heterogeneous mobile networks. In combination with common control protocols and open interfaces, this coordination can be further improved.

Related to developments at the level of hardware vendors and software developers discussed above, network operators will be able to choose among more vendors, and face increasing possibilities to develop their own software, based on open source packages such as OpenStack and OpenDaylight. This facilitates the offering of new services, compared to the current situation (and to scenario A).

[30] While energy consumption in networks may be relatively small compared to servers, some observers point out that software-defined control architectures reduces energy consumption, because of the elimination of a control plane in the network switches, less network devices, and improved utilization of network equipment. However, additional OpenFlow controllers may increase energy consumption.

Effects on competitive interaction between operators

The second type of effect on network operators pertains to competitive interaction among them (recall that above, we had abstracted from competition among networks). One can distinguish between a tendency toward concentration at the infrastructure level, and a tendency towards more competition at the services level — where 'services level' will encompass not only services to end-users, but also the provision of (virtual) infrastructure at the wholesale level and to large corporate customers. SDN enables network operators to share the physical grid, increasing economies of scale and reducing duplication in physical infrastructure (in particular the access network) — assuming that the competition authority and telecommunications regulator allow it. This translates into more cost efficiency, contributing to static efficiency. Let us take a closer look at this.

Active sharing of resources by network operators (assuming, for now, that they have incentives to offer such access) may reduce the level of service differentiation. This increases the intensity of competition at the services level. At the same time, SDN allows operators to implement independent management control over the shared infrastructure, and increases the scope for service differentiation (see below for an elaboration). Furthermore, active sharing of resources may lead to more upstream concentration and market power, given the increasing economies of scale, and hence the stronger tendency towards a (natural) monopoly at the infrastructure level. These effects may reduce the intensity of (upstream) competition. However, the possibilities to enter the market without making large investments in physical infrastructure increase. If network operators compete, at the wholesale level, to provide access, the intensity of competition at the services level may increase. The latter effect may be more prominent in mobile than in fixed, due to the (somewhat) larger number of mobile networks that compete at the wholesale level to provide access to service providers. Overall, there is at least a substantial potential for more competition (at the service level), leading to higher static efficiency. Dynamic efficiency may also increase, if entry invites innovation.

Effects on the relationship between operators, suppliers and wholesale customers

The relationship between network operators as suppliers, and various types of wholesale customers that they serve, will change. First, consider service providers that purchase network access. Related to the possibility of network sharing discussed above, SDN/NFV enables network providers to offer virtualized access to their physical infrastructure to remote parties, allowing for expansion by offering broader functionality (such as cloud access, storage and end-user device capabilities). For instance, access network providers can offer access on an on-demand basis, by using content caching and data plane processing, allowing to use network capacity in a more efficient way in response to peak usages. New business models may emerge, based on the use of time/location data used for deciding how resources should be shared. Access network providers may increase network revenues by facilitating secure and managed access to selected parts of their access networks, to other network providers, content providers, and resellers that add branding. Virtual service infrastructure providers (VSIPs) can combine virtualized infrastructures offered by different IPs into an end-to-end vertical service infrastructure. This gives them more flexibility in using access offerings. For instance, MVNOs can develop value-added services (distinct from reselling mobile operators' services), targeted at a niche market, possibly in specific sectors such as health care. Summarizing, new roles may emerge if virtual network operators lease infrastructure in order to offer added value services. This will give an impulse to dynamic efficiency. An open question is through which interface service providers will get access. Overall, service providers become more like software developers, using network access to reach their customers.

Second, consider OTT service providers. By implementing NFV and SDN, network operators can increase control over their network architecture that allows for more effective management and mining of customer data. This may enable them to make more effective use of user data, and by doing so, strengthen their position in relation to OTT service providers. NFV/SDN also allows network providers to provide more functionality to OTT providers. SDN control principles and NFV-based infrastructure allow for a fundamentally different way of building, deploying and controlling broadcast services built on top of flexible networks. This allows for dynamic and elastic delivery of high-bandwidth broadcast and media content. Thus NFV may help operators to strengthen their positions towards OTT providers, as they will have more scope to introduce innovations. Hence competition, which in the current situation leads to a lot of pressure to offer more bandwidth at lower prices, will alleviate this pressure, and allow network operators to put more weight on adding value in network services. Again there is a positive impact on dynamic efficiency.

Thus, one the one hand, OTT providers strengthen their positions in specific parts of the internet infrastructure. On the other hand, network operators get possibilities to enrich the services that they offer to end-users, which may allow them to gain position back from OTT providers. At the same time, network operators will be able to offer more functionality to OTT providers (access networks may be tuned for OTT services), so that the mutual relationships between networks and OTT providers may become more complimentary, and hence, more balanced in terms of relative bargaining positions. Hence, thanks to SDN, content providers may benefit from interfaces that allow for an improved delivery of OTT services. SDN may actually provide a framework for (more fruitful) cooperation between content providers and network operators. Consequently, static and dynamic efficiency may increase.

Thanks to NFV, service providers (OTT providers and X-play telecoms providers) can, instead of dealing with ISPs, deal with VSIPs, by leasing (optimized, service-specific) virtual service networks from them. For example, a video on demand (VoD) service provider may use a virtualized network connecting to end-users, while caching content in access network nodes. This may allow, as an example, for offering very high definition video with low latency, and lower (overall) network utilization.

Effects on end users

Regarding end-users, one may expect that they will also benefit from network virtualisation, first thanks to improved and new functionalities, and second due to cost reductions that translate into lower prices — assuming that competition is sufficiently effective. Consumers may be able to get tailor-made networks, offered by network providers or virtual network providers leasing capacity from existing networks. NFV may lead to outcomes in which end-users perceive higher QoS, assuming that the possibilities to deliver pre-agreed guarantees. Thanks to SDN, end-users may experience a smoother network experience and value from services due to improved coordination and customized (targeted for specific subscriber groups) control of different networks. Corporate customers often operate IP converged networks for voice, data and video traffic. SDN enables network operators to dynamically and quickly adapt their networks to varying QoS levels.

Overall effects

Overall, at all levels of the value chain, and in all scenarios B, C and D, static efficiency is likely to increase, relative to scenario A. The reason is that one can envisage various ways in which cost-efficiency will increase, while creating an improved (fast, flexible, more responsive to heterogeneous demands) matching of supply and demand. Also, economies of scale at the level of infrastructure (e.g. due to more effective network sharing) may reduce duplication of infrastructure, assuming that such sharing arrangements will be allowed by the competition authority. This overall impact on static efficiency is likely to be larger in scenario D than in scenarios B and C.

Scenario B is based on the assumption that SDN/NFV is implemented on proprietary, specific network equipment supplied by traditional vendors. Hence NFV/SDN does not necessarily weaken the position of hardware vendors in relation to their customers, such as network providers. DCs are, to some extent, able to increase the flexibility and cost-efficiency of their networks.

Scenarios C and D are based on the assumption that SDN/NFV is implemented on generic, white-label network equipment, supplied by a plethora of vendors. DCs are, to a larger extent than in scenario B, able to increase the flexibility and cost-efficiency of their networks. NFV/SDN increases the demand for commodity hardware, weakening the relative importance of dedicated hardware vendors. Hence NFV/SDN relatively strengthens the position of software developers and the customers of hardware vendors, such as network providers, reducing customer-lock-in. The costs of operating networks will decrease. Network operators can scale up (or scale down) network functionality, without having to commit to (or abandon) dedicated hardware.

Note in particular that in a regime of open SDN/NFV platforms (scenarios C and D), it is likely that it becomes easier for network operators to strive for openness as well. This does not necessarily mean that they will have incentives to do so, as was discussed above. Under circumstances in which network operators have incentives to provide access, scenarios C and D will exhibit an additional increase in static efficiency. Moreover, in that case one may expect more competition by alternative and virtual service providers, so that static efficiency at that level of the value chain will also increase.

Similarly, dynamic efficiency will increase in scenarios B, C and D, relative to scenario A. The reason lies in the emergence of new possibilities to invest and innovate that will become possible, at all layers of the value chain. In scenarios C and D, because interfaces are open and standardized (such as OpenFlow), SDN leads to more room for innovation by outsiders, independent of hardware vendors. In particular, there is more scope for innovation by network operators, using standard networking hardware, with less dependence of vendors of network infrastructure equipment. At the same time, with open control interfaces, network equipment vendors will have more flexibility to implement network functions and integrate their equipment into operators' networks. They can benefit from a shorter time to market. Also in scenarios C and D, assuming that network operators have incentives to provide voluntary wholesale access, one may expect more innovation by alternative and virtual service providers. The reason is that they will experience more possibilities to design services based on wholesale access to incumbents' (fixed and mobile) networks. Arguably, network operators may have incentives to keep the intensity of services competition somewhat limited, to prevent too much cannibalization with their own services, and to make sure that service operators make sufficient margins (so that there is room for margins in network operators' wholesale prices as well).

The overall impact on dynamic efficiency is likely to be larger in scenario D than in scenarios B and C. Comparing scenarios B and C is not straightforward:

- scenario B benefits from a higher adoption rate;
- scenario C exhibits a more open environment for innovation, which invites independent parties to develop and implement applications of SDN/NFV, with less interference of established players.

To conclude, both static and dynamic efficiency will increase in all scenarios B, C and D, compared to scenario A. There are differences across scenarios, though. Because there are so many factors involved, it is difficult to make clear-cut comparisons. Nevertheless, a cautious conclusion is that regarding static efficiency, the overall impact on static as well as dynamic efficiency is likely to be larger in scenario D than in scenarios B and C. This is not surprising, because (1) high adoption, by definition, leads to more matching of supply and demand, and more critical mass to support the development of process and product innovations; and (2) an open paradigm tends to be associated with more scope for adoption and innovation. Without making a large number of (possibly heroic) assumptions, it is not feasible to make an overall comparison of scenarios B and C. Arguably, policy makers should try to aim at scenario D. However, scenario analyses typically do not offer insights into specific types of public policy.

4.3.2 Efficiency driver 2: Anti-competitive practices

A priori, there is no difference in the impact of market structure developments on anti-competitive practice among the different scenarios, that is, one would need to make various and many assumptions to be able to make clear statements.

In general in markets with vertical relationships, a potential risk is vertical foreclosure, that is, exclusion that occurs when a downstream buyer is denied access to an upstream supplier (input foreclosure, or "raising rivals' costs"), or when an upstream supplier is denied access to a downstream buyer (customer foreclosure, or "reducing rivals' revenues"). Such behaviour may also come to the surface in the form of quality reductions — think of throttling — of a network input or access to downstream buyers. To assess this, the following questions are relevant: does NFV/SDN create the ability to foreclose? If yes, do firms have an incentive to engage in business conduct that forecloses competitors? In that case, is there harm to competition and consumers?

A necessary condition for a confirmation to the first question (ability to foreclose) is that a firm has significant market power. This condition may apply to established network operators with nationwide coverage, depending on the definition of the relevant market. In some market segments at the level of network operators, customer access networks may form a bottleneck — in the sense that to reach end-users, one needs to have access to the bottleneck, while duplication is not economically viable. At other levels of the value chain, the indications for such bottlenecks are much less pronounced. Therefore, we limit the discussion of anti-competitive behaviour to the level of network operators in the value chain. This is not to say that anti-competitive behaviour at other levels does not, or cannot, occur. To the contrary, at any level, there is the (at least theoretical) risk of cartel behaviour, and at certain levels, such as the OTT-level, abuse of a dominant position can be a risk.

The next step, after the ability to foreclose, concerns the incentive to do so. It is uncertain if network operators with significant market power will have an incentive to foreclose. As discussed above, the possibilities to enhance their wholesale offerings in relation to OTT providers as well as virtual operators increase. This allows network operators to strengthen their position in the value chain such that mutual benefits with wholesale customers are created. A priori, there does not seem to be a good reason to prevent network operators with SMP from developing enhanced wholesale functionality and services to OTT providers, access seekers and virtual network operators, even more so in the light of the opportunities for innovation that may be realized. Nevertheless, this does not mean that anti-competitive foreclosure will not occur. It may, and if that happens, the competition authority should be able to adequately deal with it.

Network virtualisation is likely to allow for new ways for bundling and tying of services, possibly under conditions of exclusivity, such as internet access and OTT services. In principle, competition authorities are equipped to assess if such practices are anticompetitive or not. Regarding possible breaches of net neutrality regulation, see 4.1.3.

As discussed above, the possibilities for active sharing of resources by network operators increase with NFV/SDN. Active sharing of resources may increase economies of scale at the infrastructure level, and lead to further consolidation, which may reduce competition. The competition authority will need, therefore, to continue to devote attention to potential competition problems at the level of physical access networks to end-users.

Overall effects

Anti-competitive practices tend to create immediate harm to end-users, and therefore reduce static efficiency. They may also undermine the viability of competition in general, and by doing so, reduce entry and innovation, and therefore dynamic efficiency. Overall, the risk of anti-competitive behaviour may be similar across all scenarios (at least, it is hardly possible to distinguish differences between scenarios B, C and D in this respect). Hence it is difficult to assess potential differences in static and dynamic efficiency across scenarios. Arguably, though, anti-competitive behaviour may require more attention from competition authorities in scenarios B and D, because of the higher speed of change, leading to more complexity in disentangling the various effects. A priori, we do not expect that competition authorities will have to develop specific tools and regulations, even though the current regulatory framework may not be geared towards these developments. Competition authorities will need to be prepared to diagnose new types of competition problems, irrespective of which scenario materializes. The speed of change towards a (or any, for that matter) scenario determines how fast competition authorities may be confronted with competition problems related to network virtualisation. The scenario analysis, however, makes no statements about the speed of change.

4.3.3 Efficiency driver 3: Public policy

As with regard to anti-competitive practices (4.1.2), a priori there is no difference among the scenarios in the impact of market structure developments on public policy and public interests.

The purpose of net neutrality regulation is to safeguard the access of end-users to any content and services he or she wishes to 'consume', unencumbered by gatekeepers. It is difficult to assess if NFV/SDN will affect the effectiveness of current legislation regarding net neutrality. As discussed in chapter 3, ISPs have the technical possibility to provide specialised network access for services that have specific requirements regarding QoS. Virtualisation enables the separation of applications and internet traffic, and allows for the provisioning of guaranteed bandwidth. Such practices are currently not allowed (except under specific circumstances), insofar as legislation regarding net neutrality applies to internet access services. New technologies may create unforeseen possibilities (see section 3.3). However, ISPs can use SDN/NFV to provision services over their existing access network, separate from the internet. This allows them to provide services using favourable network conditions, but without being in breach of net neutrality legislation (see section 3.3). At least in theory, this may lead to additional barriers to entry for service providers. Hence, the legislator and regulators (at the European and national level) may wish to clarify the notion of network neutrality in relationship to network virtualisation, addressing issues such as the definition of internet access. For instance, is a port on a modem at consumers' premises, that is dedicated to an OTT service, part of an internet access service or not?

Policy makers can –in principle- facilitate and support, and possibly mandate, the development and implementation of open standards (somewhat similar to European policy regarding mobile telephony standards). Note, however, that there is no obligation for operators to use an open standard. Such policies could (again: in principle) enhance the prospects of scenarios C and D, and are likely to give an impulse to dynamic efficiency.

An inherent characteristic of SDN is that the interfaces are open. Nevertheless, hardware vendors sometimes label 'closed products' as based on SDN. Note that OpenFlow (the controller’s southbound interface) an open standard is. The northbound interfaces should also be open — but they are still in development. The most important incentive to develop open interfaces is that it allows network operators to exercise more control, and gives them the ability to add new features and services. Large OTT providers already apply this principle to optimally use networks for specific services. ISPs view openness as a requirement to strengthen their position towards OTT players (who tend to have much larger market capitalisations than many national ISPs). For ISPs, it's a matter of survival. Hence there may not be a need for public intervention in this respect. Having said that, the extent to which network infrastructure operators are willing to provide interfaces to SDN and NFV functionality to third parties is a different matter. Policy makers can — in principle — facilitate and support, and possibly mandate, the development and implementation of open standards (somewhat similar to European policy regarding mobile telephony standards). Note, however, that there is no obligation for operators to use an open standard. Such policies could (again: in principle) enhance the prospects of scenarios C and D, and are likely to give an impulse to dynamic efficiency. [31]

As the regulatory framework for telecommunications regulation undergoes revisions on a regular basis, it is uncertain to what extent it will apply to active network sharing. This is even more so as it will take many years, and substantial investments, for network virtualisation to become the new paradigm in existing networks. If competition turns out to be insufficiently effective, competition authorities may want to consider various options for intervention, including access obligations. Note that providing access affects the nature of competition, and therefore it is not obvious that network operators will have incentives to provide such access. Nevertheless, if they experience more infrastructure competition, they will have stronger incentives for openness, which for instance materializes in competition at the wholesale level to provide access. To see this, note that infrastructure competition induces network operators to face a pressure that if they do not provide access, wholesale customers may turn to a competing network. Hence, providing access may come at the cost of increasing competition at the retail level, but if operators do not want to lose wholesale revenues to other networks, they will nevertheless choose to provide wholesale access (cf. a prisoners' dilemma).[32]

If virtual appliances run in data centres not owned by network operators, the introduction of new network elements (e.g. orchestrators) may create new vulnerabilities. [21] A similar observation applies to network virtualisation, as it gives rise to new types of security

[31] Note that whereas a scenario analysis as carried out in this study does not lend itself for specific policy suggestions, here we have run into an exception. The degree of openness, a fundamental uncertainty that defines one of the two axes of our scenarios, can be forced into a certain direction by simply mandating open standards and open networks. Having said that, note that such an intervention would require a large, coordinated effort at the European level and beyond. Hence, there are limits to trying to steer the market towards an 'open' scenario, just as there are limits to stimulating the adoption of NFV/SDN (the other axis/fundamental uncertainty).

[32] The economic literature shows that in general (hence exceptions are possible), infrastructure competition tends to create downward pressure on wholesale tariffs for network access, and that foreclosure is unlikely. See [8], [33], [10] and [9].

vulnerabilities, such as the risk that an attack against a physical network in a virtualized environment will affect all virtual networks that are hosed on it. [13] Also, the sharing of networking and storage increases mutual dependence, which may also lead to more vulnerability. Using software-based components from different vendors increases the complexity of integration, and hence the risk of security threats. Effectively, the attack surface of a network becomes larger. This is not to say that NFV/SDN form an intrinsic threat to security. The point is that they will raise new issues, to be addressed at the state of system design and the development of virtualisation environments. In particular, the design and management of security may become easier, due to the possibility of specific, virtualized firewalls, that can be updated on a regular basis from a central level. The abstractions provided by SDN and NFV in theory make it possible to swap hardware from one vendor with hardware from another. If a vendor is not trusted (e.g. a backdoor is suspected) this provides a relatively painless way to fix the security issue.

In the mobile market, operators have been providing voluntary access to MVNOs for various reasons, such as the efficient use of spare network capacity by having MVNOs aim at niche markets beyond the reach of network operators. It is uncertain if a similar mechanism will become active in fixed telecommunications, although one may imagine that specialised virtual operators may be more effective in developing applications for specific sectors (such as health care) than network operators. By using SDN/NFV, network operators may implement advanced QoS policies that allow for selling spare capacity to virtual operators without cannibalizing the core (retail) business of networks. It remains to be seen how much control over the network virtual operators will get, compared to current wholesale access services.

Overall, regarding access obligations, for now it is important to monitor SDN/NFV developments, in order to better be able to anticipate the regulatory implications. Network virtualisation is still in an early stage of development and implementation, and it is important to prevent the distortion of innovation and adoption, in order to avoid the harming of dynamic efficiency.

If virtual appliances run in data centres not owned by network operators, the introduction of new network elements (e.g. orchestrators) may create new vulnerabilities. [21] A similar observation applies to network virtualisation, as it gives rise to new types of security vulnerabilities, such as the risk that an attack against a physical network in a virtualized environment will affect all virtual networks that are hosed on it. [13] Also, the sharing of networking and storage increases mutual dependence, which may also lead to more vulnerability. Using software-based components from different vendors increases the complexity of integration, and hence the risk of security threats. Effectively, the attack surface of a network becomes larger. This is not to say that NFV/SDN form an intrinsic threat to security. The point is that they will raise new issues, to be addressed at the state of system design and the development of virtualisation environments. This may make networks more resilient to threats. In particular, the design and management of security may become easier, due to the possibility of specific, virtualized firewalls, which can be updated on a regular basis from a central level. The abstractions provided by SDN and NFV in theory make it possible to swap hardware from one vendor with hardware from another. If a vendor is not trusted (e.g. a backdoor is suspected) this provides a relatively painless way to fix the security issue. Note, furthermore, that SDN implies that more software is used, which introduces different types of risks (such as errors in code).

To the extent that functions that are virtualized involve personally identifiable information that is transferred to the cloud, new challenges arise. [28] When functions are distributed, it becomes harder to know where data is located and how has access to it. In the case of third-party clouds, users, network providers and service providers do not have access to the physical security system of data centres. Network providers and service providers may specify their privacy and security requirements in a contract, but this does not guarantee that they will be fulfilled, let alone that such contracts foresee all possible risks. Using certain forms of SDN may make it easier to gather metadata. A growing deployment of NFV and an increase of functions that are virtualized, will attract new threats to security and privacy, including threats on data interception. Nevertheless, network virtualisation does not form an inherent threat to privacy, as new issues will be addressed at the design stage on an ongoing basis.

Overall effects

All these public policy issues relate to static and dynamic efficiency. Security and privacy can be seen as public interest concerns, which constrain the policy goal of maximizing static and dynamic efficiency. In the light of the growing dependence on electronic communications services and networks, security and privacy have become much more prominent — a trend that can be expected to continue. Arguably, they may require more attention from policy makers in scenarios B and D, because of the higher speed of change, leading to more complexity, more intrusion risks, and hence more challenges for design and development. To the extent that an open paradigm for network virtualisation creates more transparent technologies, security risks may be somewhat less pronounced in scenarios C and D, relative to A and B.

4.3.4 External effects for the economy and society as a whole

Electronic infrastructure and communications, and more generally, ICT, are known for their positive external effects for the economy and society as a whole. If anything, these spill overs will increase due to any development that renders these 'general-purpose technologies' more effective. Network virtualisation is such a development, as has become clear from the discussions above. As almost all sectors use ICT, all businesses and organisations will, at some point, experience the technological changes, even though they may remain hidden from them for a while.

By definition of the scenarios, it is likely that scenarios B and D, which are characterized by high adoption of network virtualisation technologies, will have a larger positive impact on the positive externalities than scenarios A and C. To the extent that an open paradigm leads to more dynamic efficiency than a closed paradigm, scenarios C and D will have a bigger impact on the spill overs than scenarios A and B. Thus, scenario D seems to be preferable not only from the (somewhat narrow) viewpoint of dynamic efficiency, but also from a broader welfare perspective.

4.4 Overview

Overall, the exploration of the efficiency drivers, and (to the extent that this was feasible) the impact in the different scenarios, showed that network virtualisation is beneficial for static and dynamic efficiency, and will strengthen the positive externalities of ICT for the economy and society as a whole. Nevertheless, the impact on specific public interests, for instance related to cybersecurity, are uncertain, as new risks may come to the surface, while at the same time, networks and applications may become more resilient to threats. By construction, scenario D (open paradigm, high adoption) sketches the most attractive perspective for welfare, both in the short and in the long run. The scenario analysis, however, is not able to establish the likelihood that a given scenario materializes. Similarly, the analysis is not suited to identify policy proposals that make scenario D more likely to come about. The purpose of the exploration is different, namely to identify the fundamental uncertainties, and to understand what the impact of network virtualisation may be, depending on conditions and circumstances that one may currently foresee, and that may vary across the scenarios.

5 Conclusion

5.1 The road ahead for SDN and NFV

Network virtualisation stimulates innovation by, from a technical point of view, enabling diverse network architectures to cohabit on a shared physical infrastructure. To make this possible, network elements need to be 'programmable' - software-defined networking (SDN) makes the control plane programmable, while network function virtualization (NFV) does the same for the data plane. By doing so, network architecture complexity becomes lower (e.g. easier network configuration via a centralized SDN controller), network-related expenditures are reduced (e.g. NFV promotes the usage of programmable general-purpose hardware), and it becomes less demanding to innovate (e.g. network operators become less dependent on standards development organisations and vendors to introduce new features).

It is important to realise that network virtualisation is an evolutionary and open-ended process, in line with similar developments in the wider ICT industry, related to virtualisation of storage and computing capacity. SDN and NFV jointly form the next logical step in this broader development. For many network operators (i.e., ISPs), network virtualization, SDN and NFV are (relatively speaking) still somewhat out of sight. Others are early adopters, and are already reaping benefits from these technologies (i.e. several OTT providers).

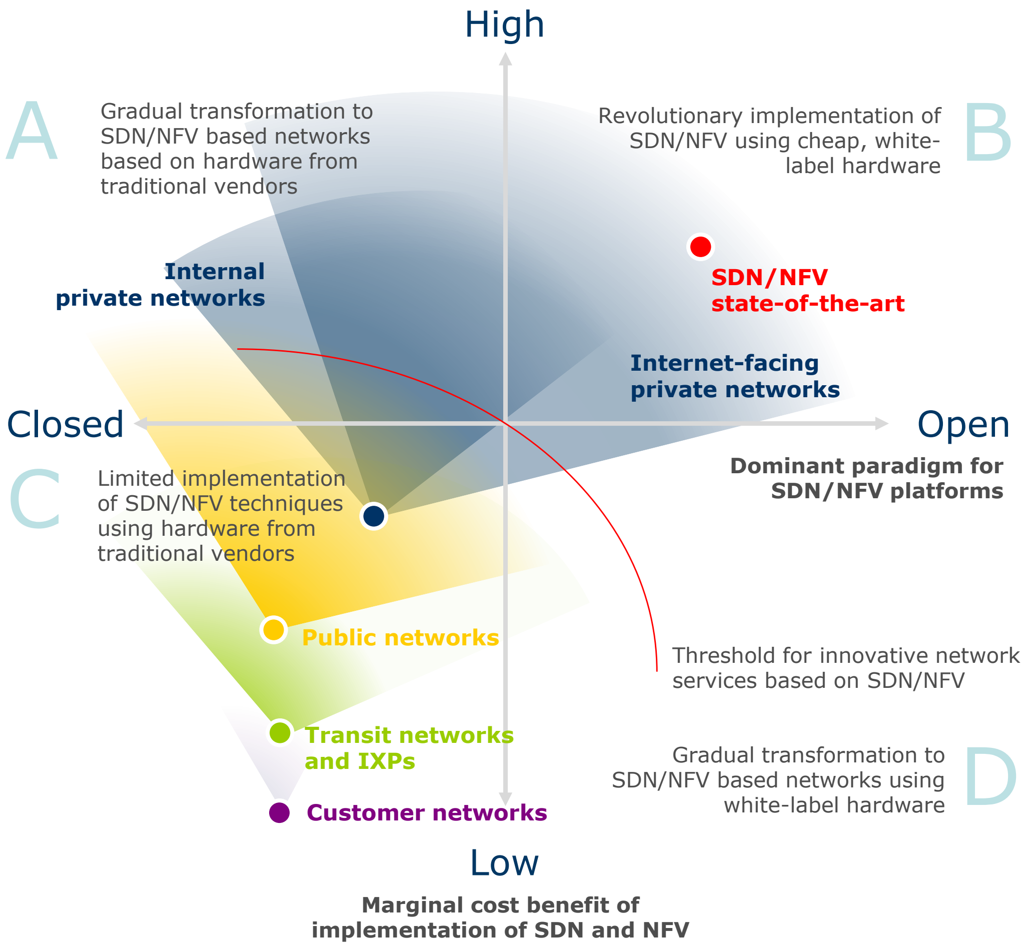

High
A
Gradual transformation to SDN/NFV based networks based on hardware from traditional vendors
B
Revolutionary implementation of SDN/NFV using cheap, white-label hardware
SDN/NFV state-of-the-art
Internal private networks
Internet-facing private networks
Closed
Open
Dominant paradigm for SDN/NFV platforms
C
Limited implementation of SDN/NFV techniques using hardware from traditional vendors
Public networks
Threshold for innovative network services based on SDN/NFV
Transit networks and IXPs
Gradual transformation to SDN/NFV based networks using white-label hardware
D
Customer networks
Low
Marginal cost benefit of implementation of SDN and NFV

Figure 8 Outlook for adoption of SDN and NFV in different playgrounds, and in different scenarios

In this study, we have analysed four scenarios for adoption of SDN and NFV technologies (one of which corresponds to the current situation), distinguishing different possible routes for market developments, and within them, different 'playgrounds' where these developments may materialise. The two primary drivers for the scenarios are (1) the *degree of openness* of the underlying technology stack as well as the infrastructures on which SDN and NFV are deployed, and (2) the *level of adoption* of SDN/NFV technologies, following primarily from the advantage provided by them.

Many of the new propositions made possible by SDN and NFV require open access to the SDN and NFV technology stack, as well as the ability to create virtual networks on top of infrastructure owned by other players on the market (Network as a Service). A higher degree of openness is expected to lead to a higher degree of innovation in the area of SDN and NFV, which primarily benefits private networks as well as (certain parts of) ISP networks, where the benefits of SDN and NFV are the greatest.

For SDN and NFV to become adopted more widely, they need to provide a significant technical advantage, or alternatively, an organisational advantage. From a technical perspective, many of the features made possible by SDN and NFV can already be realised today, by using existing technologies. However, the novel aspect lies in the way in which functionalities can be developed and implemented — SDN and NFV primarily improve efficiency, cost-effectiveness and manageability of existing networks.

Figure 8 shows what can be expected in different playgrounds in the different scenarios in the time horizon of roughly five years:

- On *end-user networks*, the opportunities for SDN and NFV are relatively limited in scope (compared to the possibilities at higher network levels) and highly specific.

- On *transit networks*, performance requirements have historically dictated the use of leading edge proprietary networking equipment. The opportunities for applying NFV are limited, as transit operators typically offer little networking services beyond traffic transportation at the lower network layers. Opportunities for SDN and NFV are hence relatively limited.

- On *public ISP networks*, SDN provides opportunities for managing the (large) complexity associated with heterogeneous customers and traffic. Additionally, SDN can support even more complex configurations than currently exist, such as scenarios where services are delivered to end-users over separate, virtual channels.

  For ISPs, NFV is primarily of interest when applied at the edges of the network, where ISPs may offer new services related to network security and content delivery. The virtualisation of CPEs (modems at customer premises) is of particular interest, as the CPE is typically installed at a strategic location for provisioning latency-sensitive services. While virtualized CPEs could also be used to facilitate consumer switching between ISPs, the ISPs currently own the CPEs, and will not have an incentive to implement it if there is no (regulatory) obligation to do so.

- On *internet-facing private networks*, implementation of SDN and NFV appears to be a no-brainer for many organisations. In such deployments, SDN and NFV remain inside the boundaries of a data centre, where they allow for a more efficient usage of resources, lower costs (if open, general purpose hardware can meet the demands) and more efficient management.

- On *internal private networks*, SDN and NFV could provide similar benefits as those for internet-facing private networks, depending on the specific type of network.

For existing networks, the openness of technology and infrastructure is a less important factor with respect to the adoption of SDN and NFV. In transit and IXP networks, stringent performance requirements have lead the operators of these networks to use proprietary, leading edge technology. In addition to technical requirements, these operators require the support, training and guarantees of hardware vendors.

5.2 Answering the research questions

5.2.1 Business models and applications

Which new (business) models are enabled by network virtualisation for market parties as well as society? Which applications are currently foreseen, and what use cases and business cases are regarded by organisations who have been adopting virtualisation technologies as the most promising in the short term?

While SDN and NFV are sometimes presented as being revolutionary networking technologies, we find that these technologies are rather the next logical step in an evolutionary, albeit open-ended, process of virtualisation, which follows developments in the wider computer industry.

From a technical perspective, the functionality offered by NFV/SDN is already possible by using currently existing technologies. In fact, many new business models for NFV and SDN simply run on top of existing networks, and can already be deployed today, as illustrated by services that allow to create virtual layer 2 networks on top of the internet. [40]

While we do not foresee revolutionary business models based on the technical merits of SDN and NFV, we do expect evolutionary introduction and modification of business models based on the incremental improvements in organisational efficiency provided by SDN and NFV:

- SDN and NFV cut down the complexity of network management in heterogeneous environments. This enables business models that require complex network topologies, such as edge computing/fog networks and those using virtual CPEs. These topologies can become highly important in future networks, given developments in automotive and virtual reality which require low-latency communication.

- SDN can provide opportunities for business models that require certain forms of QoS (quality-of-service) that cannot currently be provided over-the-top. Use cases are for instance those related to utilities (smart metering) and healthcare.

- SDN and NFV make it easier for infrastructure owners to open up their networks to third parties at lower levels. However, network infrastructure owners likely do not have the incentive to provide said access. The primary driver for this application of SDN and NFV technology is therefore not the technology itself, but competitive pressure (at the wholesale level) and commercial opportunities (selling spare capacity to virtual operators active in niche markets) to provide such access, and if necessary and desirable from a welfare perspective, regulation.

5.2.2 The competitive landscape

How will virtualisation change the competitive landscape of ISPs, network vendors and service providers, and how will it change their stance towards standardisation of, research & development (R&D) on, and deployment of SDN and NFV technologies?

Hardware vendors will have to adapt to an evolving customer demand for SDN/NFV-based products. Some of them may benefit from selling specialised software that runs on general purpose hardware. Similarly, some software developers may specialise in specific functionalities. Their possibilities to enter the market for telecommunications services increase, because their expertise increasingly becomes a core element of the infrastructure. Virtualisation allows DCs to meet more stringent requirements on their networks, following from an increasing heterogeneity of services.

Network operators will increasingly implement remote management of physical networks and end-to-end service management, as well as flexible scaling of capabilities. One may anticipate a tendency toward concentration at the infrastructure level, due to economies of scale and more effective network sharing, and towards more competition at the services level, which encompasses the provision of (virtual) infrastructure at the wholesale level and to large corporate customers. Virtual network operators may introduce new business models in order to offer added value services in niche markets.

OTT providers will be able to strengthen their positions in specific parts of the internet infrastructure, while network operators may enrich their end-user services, allowing them to gain position back from OTT providers. At the same time, the mutual relationships between networks and OTT providers may become more complimentary, which is beneficial to both types of suppliers.

At all levels of the value chain, static efficiency is likely increase in scenarios B, C and D, relative to scenario A, due to various ways in which cost-efficiency will increase, while creating an improved matching of supply and demand. Also, economies of scale at the level of infrastructure may reduce duplication of infrastructure, if such sharing arrangements are allowed. In the open scenarios C and D, it may become easier for network operators to strive for openness. Conditional on network operators having incentives to provide access, these scenarios will exhibit an additional increase in static efficiency. One may then also expect more competition by alternative and virtual service providers, so that static efficiency at that level of the value chain will also increase.

Dynamic efficiency is likely increase as well in scenarios B, C and D, relative to scenario A, because of new possibilities to invest and innovate at all layers of the value chain. In the open scenarios C and D, there will be more room for independent innovation by outsiders, including network operators. At the same time, network equipment vendors will have more flexibility to implement network functions and integrate their equipment into operators' networks. Also, if network operators have incentives to provide voluntary wholesale access, one may expect more innovation by alternative and virtual service providers. Thus these scenarios tend to lead to higher dynamic efficiency.

The overall impact on static and dynamic efficiency is likely to be larger in scenario D than in scenarios B and C, because (1) high adoption leads to more matching of supply and demand, and more critical mass to support process and product innovations; and (2) an open paradigm tends to be associated with more scope for adoption and innovation. A priori, there does not seem to be a tendency towards anti-competitive behaviour, although it seems wise that market authorities monitor market developments. Anti-competitive behaviour may require more attention in scenarios B and D, because of the higher speed of change, leading to more complexity.

Although SDN implies that interfaces are open, hardware vendors sometimes label 'closed products' as based on SDN. The most important incentive to develop open interfaces is that it allows network operators to exercise more control, and gives them the ability to add new features and services. Large OTT providers already apply this principle, while ISPs view openness as a requirement to strengthen their position towards OTT players.

5.2.3 Points of control

What are the points of control in virtualised networks, and which market parties will have access to these in the future?

The central point of control in networks based on SDN and NFV technologies is the orchestration layer. From this control point, a software defined network can be configured, and virtualised network functions can be created in the infrastructure. The orchestration layer is usually operated by the owner of the physical network infrastructure.

Operators of public or transit networks could make accessible an interface for others to define virtual networks and network functions on top of the infrastructure, and quickly alter their characteristics (providing *Networks-as-a-Service*). ISPs could also offer an interface using which customers can enable various filters and QoS preferences for their internet connection.

Having access to virtualisation functions is crucial in realising the so-called end vision for SDN and NFV, where a virtual infrastructure provider can ‘mix and match’ various types of physical connectivity in order to create a network service. The provided interface could either be based on an open standard, or it could be more closed and specific. Standardised SDN and NFV related interfaces (such as the OpenFlow protocol) are evolving from one version to the next. The trend is however to create a wide interface, in the sense that the interface provides a great level of control over various aspects of the network. Specific implementations, such as vendor implementations, of the interface could however be more restricted (narrow). In order to be able to deploy SDN/NFV based solutions to consumers and SMEs, access to particular control points on the access network may be required.

In private networks, the orchestration layer is expected to develop to also include computing and storage resources. It is likely that this type of control point will be closed and vendor/platform-specific.

Direct (physical or logical) access to networking equipment will stay relevant for the years to come, due to the fact that not all functionality is available when an abstraction layer is used. Also, for debugging purposes, direct access remains relevant.

5.2.4 Regulated network access

Which types of regulated network access are needed to allow for effective competition and market entry of alternative service providers, who do not own an access network?

It is too early to answer this question, as SDN/NFV is still immature in its development as well as adoption. These developments should be monitored, in order to anticipate the potential need (if any) for adapting or maintaining access obligations.

SDN and NFV technologies may aid operators of public networks to comply with current network access regulations in novel ways. In addition, the technology may provide additional opportunities for providing network access at low levels.

In principle, public network operators may use SDN and NFV technology to provide access to third parties (this does not say that they will always have incentives to do so on a voluntary basis). A network operator could for instance provide layer 2 access to a specific service provider, in order to offer a particular service with certain guarantees regarding quality of service. This would be similar to the current situation, where network operators are offering services such as television and telephony over the same access infrastructure, but using different virtually defined communications channels, which specific quality of service attributes. Note that current access obligations do not include specific prescriptions regarding SDN and NFV. Nevertheless, providing access based on NFV/SDN could be subject to current legislation, depending on the situation.

Network infrastructure owners could, in principle, also provide third parties with the ability to create virtual networks on top of their infrastructure (also here: they may not always have the incentives to offer such access). These could be used to provide connectivity for applications that currently have reverted to other types of connectivity, as general internet connectivity could not provide the desired quality of service. An example are smart meters – in many cases, the smart meter will be right next to an internet modem, so it seems logical to use that connection to transmit the required data. However, in the Netherlands, all smart meters are connected wirelessly, either using LTE, or using a CDMA-based network specifically deployed for this purpose. Whether providing access that allows for setting up virtual networks is subject to current access obligations, will depend on the situation at hand (e.g., if such access is used to deliver services that are included in the Commission Recommendation on relevant product and service markets within electronic communications).

Whether (adaptations of) access regulation will be necessary, will depend on various factors, such as technological developments regarding SDN/NFV and the nature of competition between network operators (note that this does not imply that such obligations are socially optimal, which will depend on their impact on static and dynamic efficiency). As illustrated by the fact that at present, mobile operators voluntarily offer access to MVNOs, competition between network operators may incite them to offer wholesale access without obligations to do so. Also, the potential for complementarity between networks and applications may create new business cases for mutually beneficial vertical agreements.

For now, it is important to first of all monitor market and technological developments. Physical access to networking equipment as well as access to lower layers (layer 2 specifically) appear to remain relevant for the next coming years as SDN and NFV mature. While this may not require access regulation to be changed just yet, we advise to investigate whether policy can be changed such that(in the future) access to (certain parts of) a network orchestration layer may be regulated.

5.2.5 Net neutrality

What influence can virtualisation have on net neutrality, and how should this be monitored from a regulatory point of view? In particular, what role can SDN and NFV play in improving quality of service (QoS) aspects?

Network virtualisation is closely linked to net neutrality, but it is not yet clear how this relationship will develop. Net neutrality applies to any connection that provides connectivity to the internet. Net neutrality regulation requires that all traffic on a connection be treated similarly – service providers may not prioritise one type of traffic over the other, except in very specific cases where it is needed to prevent network congestion. However, even with this exception it may not be possible to adequately identify traffic that requires special treatment to prevent network congestion, as net neutrality regulation does not allow the usage of deep packet inspection (DPI) to classify traffic[33] (and define traffic flows). This does not only make it impossible/hard to prioritize traffic from certain parties (e.g. an OTT SPs traffic is prioritized over other traffic) but also to prioritize certain types of traffic independent of the originating party (e.g. live video traffic cannot be prioritized over less delay sensitive applications such as video streaming). This may prevent the Internet from becoming a suitable platform for certain service types.

Virtualisation allows the creation of virtually separate connections to consumers and SMEs that are not necessarily considered as part of 'the internet' but do provide access to a service. Hence, network operators may argue that net neutrality regulation does not apply to virtual networks. As such network operators could argue that a service can be offered over a separate virtual connection with service-specific quality-of-service (QoS) without breaking net neutrality regulation. While that might, formally, not lead to a conflict with legislation, one may argue that it would undermine the intent of the law, which aims at openness to stimulate, among others, innovation by independent firms. Therefore, keeping the underlying idea and motivation of net neutrality in mind, one could argue that the rules should apply, more broadly, to network connections in general — in particular when network providers have substantial market power in markets for access to end-users (it is beyond the scope of this study to make a legal assessment of the need for an adaptation of current legislation). The desirability of such an interpretation or extension warrants further study, pertaining to the nature of dynamic competition, including the incentives and possibilities to invest and innovate by network operators as well as independent developers.

Note that from an operational point of view, without SDN and NFV, deploying such virtual networks would be highly complex: an ISP would have to manage the configuration of a large number of VLANs for different service providers and subscribers. SDN and NFV allow for the flexible creation of private networks at layer 2. While this is currently also possible (and done) using (over-the-top) VPN technologies, the virtual networks at layer 2 are much more efficient and reliable. SDN and NFV can be used to set various QoS parameters (e.g. guaranteed latency/jitter). While QoS guarantees could also be configured at the IP level, this would conflict with net neutrality legislation.

Without further clarification with regard to regulation of virtual networks, network operators will need to turn to the market authority in order to decide which services are provisioned 'outside the internet' and which remain on the over-the-top connection. If the decision remains at the discretion of the network operators, it might, at least in theory, lead to additional barriers to entry for service providers. To guarantee the aim of net neutrality whilst at the same time having market parties and society reap the benefits of SDN and NFV, a (continued) dialogue between regulators and market parties is advisable.

5.2.6 Security and privacy

How will SDN and NFV technology influence the security and privacy of network communication? Is there a need to change or update regulation regarding privacy and/or security requirements?

Overall, new threats to security, as well as opportunities to make networks more resilient, may arise.

SDN and NFV enable functional separation of networks without physical separation. This may provide both benefits and well as risks regarding network security: while networks can be

[33] Most traffic cannot be classified solely based on header information as such a peek into the payload of a packet may be required.

separated more easily, the risk that the virtual barrier gets compromised is higher than in the case of a physical barrier, where guarding access is more straightforward. However, as the industry is moving to end-to-end encrypted connections for most applications, security does not seem to be a main argument for or against implementing SDN and NFV.

The logical centralisation of control in SDN creates an attractive target for attackers. Obtaining control over the orchestration layer of a network would put an attacker in control of network traffic as well. It would theoretically be possible to install taps or to perform man-in-the-middle attacks by controlling the orchestration layer. Using certain forms of SDN, it would also be easier to gather metadata.

While centralisation of control reduces the need to configure hardware and software network components individually, there remains a need for direct access for troubleshooting and more complex configuration. Effectively, the attack surface of a network becomes larger.

Centralisation of control provides a way to easily audit the security and traffic policies in a network for compliance with rules and regulations (e.g. on privacy, net neutrality). This of course assuming that the central policies are correctly translated to configurations of the individual components.

Finally, the abstractions provided by SDN and NFV in theory make it possible to swap hardware from one vendor with hardware from another. If a vendor is not trusted (e.g. a backdoor is suspected) this provides a relatively painless way to fix the (suspected) security issue.

References

[1] ADVA (2016). *Mitigating Security Risk in Practical vCPE Solutions* [linuxfoundation.org]

[2] Aghion, P. N. Bloom, R. Blundell, R. Griffith and P. Howitt (2005), "Competition and Innovation: an Inverted-U Relationship", *Quarterly Journal of Economics* 120 (2), p. 701-728.

[3] Amin, F. and S. Feizi (2014), "Big Data Strategy for Telco: Network Transformation", *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering* 8(2), p. 457-460.

[4] Araújo, J. (2016). *Building and scaling the Fastly network, part 1: Fighting the FIB.* [fastly.com]

[5] Barroso, D. (2016). *SDN Internet Router, part 1.* [spotify.com]

[6] Barroso, L. A., Clidaras, J., & Hölzle, U. (2013). *The datacenter as a computer: An introduction to the design of warehouse-scale machines.* Synthesis lectures on computer architecture, 8(3), 1-154.

[7] Bennett, M. , P. de Bijl and M. Canoy (2001), "Future Policy in Telecommunications: An Analytical Framework", *CPB Document 005,* CPB Netherlands Bureau for Economic Policy Analysis, The Hague.

[8] Bijlsma, M. en M. van Dijk (2008), "Netwerkconcurrentie en toegangsregulering in de telecommunicatiesector", *TPEdigitaal* 2(1), p. 62-78.

[9] Bourreau, M., Hombert, J., Pouyet, J. en Schutz, N. (2011), "Upstream Competition between Vertically Integrated Firms*", Journal of Industrial Economics* 59(4), p. 677-713.

[10] Brito, D. en P. Pereira (2010), "Access to Bottleneck Inputs under Oligopoly: A Prisoners' Dilemma?", *Southern Economic Journal* 76(3), p. 660-677.

[11] Chappell, C. (2013), "Unlocking Network Value: Service Innovation in the Era of SDN", *White Paper on behalf of Ciena, Cisco and Skyfire*, Heavy Reading, June.

[12] Chen, T., M. Matinmikko, X. Chen, and P. Ahokangas (2015), "Software Defined Mobile Networks: Concept, survey, and Research Directions", *IEEE Communications Magazine*, November, p. 126-133.

[13] Chowdhury, N. M. K., & Boutaba, R. (2010). *A survey of network virtualisation.Computer Networks*, 54(5), 862-876.

[14] Davy, S., A. Miron, P.M. Neves, J. Famaey, M. Dramitinos, S. Latré, J. Serrat, J.L. Gorricho, and E. Goshen (2014), "Challenges to support Edge-as-a-Service", *IEEE Communications Magazine* 52(1), p. 132-139.

[15] De Bijl, P.W.J. (2004), *"Competition, innovation and future-proof policy*", report commissioned by OPTA, TILEC (Tilburg Law and Economics Center), Tilburg University.

[16] Dialogic (2014). *De impact van ICT op de Nederlandse economie.* In opdracht van: Ministerie van Economische Zaken. [dialogic.nl]

[17] Ellerton, J., A. Lord, P. Gunning, K. Farrow, P. Wright, D. King, and D. Hutchison (2015), *"Prospects for Software Defined Networking and Network Function Virtualisation in Media and Broadcast",* Written for presentation at the SMPTE 2015 Annual Technical Conference & Exhibition.

[18] Ericsson (2014). *Virtual CPE and Software Defined Networking.* [ericsson.com]

[19] ETSI, "Network Functions Virtualisation (NFV); Architectural Framework" Okt. 2013

[20] ETSI, "Network Functions Virtualisation (NFV); Use Cases" Okt. 2013

[21] Han, B., V. Gopalakrishnan, L. Ji, and S. Lee (2015), "Network Function Virtualisation: Challenges and Opportunities for Innovations", IEEE Communications Magazine 53(2), p. 90-97.

[22] Kocsis, V., and P.W.J. de Bijl (2007), "Network neutrality and the nature of competition between network operators*", Journal of International Economics and Economic Policy* 4 (2), p, 159-184.

[23] Kreutz, D., Ramos, F. M., Esteves Verissimo, P., Esteve Rothenberg, C., Azodolmolky, S., & Uhlig, S. (2015). *"Software-defined networking: A comprehensive survey."* Proceedings of the IEEE, 103(1), 14-76.

[24] Labovitz, Craig, et al. "*Internet inter-domain traffic.*" ACM SIGCOMM Computer Communication Review 41.4 (2011): 75-86.

[25] Latré, S., J. Famaey, F. De Turck, and P. Demeester (2014), "The fluid internet: service-centric management of a virtualized future internet", *IEEE Communications Magazine* 52(1), p. 140-148.

[26] Liebenau, J., Karrberg, P., & Elaluf-Calderwood, S. "A critical analysis of the effects of internet traffic on business models of telecom operators."

[27] McKeown et al. (2008). *OpenFlow: enabling innovation in campus networks.* [acm.org]

[28] Mijumbi et al. (2015). Network Function Virtualization: State-of-the-art and Research Challenges*. IEEE Communications Surveys and Tutorials. Accepted for Publication (September 2015).*[maps.upc.edu]

[29] Naudts, B., M. Kind, F.-J. Westphal, S. Verbrugge, D. Colle, and M. Pickavet (2012), "Techno-economic analysis of software defined networking as architecture for the virtualisation of a mobile network", *2012 European Workshop on Software Defined Networking, IEEE.*

[30] Naudts, B., W. Tavernier, S. Verbrugge, D. Colle and M. Pickavet (2016), "Deploying SDN and NFV at the Speed of innovation: Toward a new bond between standards development organizations, industry fora, and open-source software projects", IEEE Communications Magazine - Communications Standards Supplement, p. 46-53

[31] OpenDaylight (2016). *OpenDaylight: open source SDN platform.* [opendaylight.org]

[32] Open Networking Foundation (2012), "Software-Defined Networking: The New Norm for Networks", *ONF White Paper*, April 13.

[33] Ordover J. en G. Shaffer (2007), "Wholesale access in multi-firm markets: When is it profitable to supply a competitor?", International Journal of Industrial Organization 25, p. 1026-1045.

[34] Pentikousis, K., Wang, Y., & Hu, W. (2013). Mobileflow: Toward software-defined mobile networks. Communications Magazine, IEEE, 51(7), 44-53.

[35] Regulation (EU) 2015/2120 of the European Parliament and of the council. [europa.eu]

[36] Rijksoverheid (2016). *Telecommunicatiewet, artikel 7.4a.* [overheid.nl]

[37] Ryan (2010) *A history of the Internet and the Digital Future*

[38] Sharma, S. "Towards high quality and flexible future Internet architectures" PhD dissertation

[39] Sherry, J., Hasan, S., Scott, C., Krishnamurthy, A., Ratnasamy, S., & Sekar, V. (2012). "*Making middleboxes someone else's problem: network processing as a cloud service.* " ACM SIGCOMM Computer Communication Review, 42(4), 13-24.

[40] Telfort (2016). *Invloed van televisie op internetsnelheid* [telfort.nl]

[41] Utility Connect (2016). *Ons netwerk*. [utilityconnect.nl]

[42] Ziggo (2016). *WifiSpots*. [ziggo.nl]

[43] Zerotier (2016). *Technical FAQ*. [zerotier.com]

Appendix 1. List of interviewees

One-on-one sessions

Organisation	Name(s)	Function title(s)
AMS-IX	Henk Steenman Bastiaan Goslings	CTO Governance and Policy Officer
Cap Gemini	Wim van der Bijl	Network & security architect
Fastly	David Barroso	Network systems engineer
Google	Edo Haveman	Public policy manager
Huawei	Jurjen Veldhuizen	Senior Marketing Manager
Nokia	Andre van Buiten	Regional Sales Director
SURF	Migiel de Vos Gerben Malenstein	Manager Network Services

Workshop session

Organisation	Name(s)	Function title(s)
TNO	Arjen Holtzer	Consultant – innovator network technology
ACM	Cees-Jeroen Bes	Technical advisor
Huawei	Jurjen Veldhuizen	Senior marketing manager
SURF	Migiel de Vos	Network, innovation & product development
Deutsche Telekom	Richard Marijs	Manager network economics

Contact:

Dialogic
Hooghiemstraplein 33-36
3514 AX Utrecht
The Netherlands

Tel. +31 30 215 05 80
Fax +31 30 215 05 95
dialogic.nl/en