



**Project No. 688503**  
**NEWTON – Networked Labs for Training in Sciences and  
Technologies for Information and Communication**

# Networking

Authors: Podhradský Pavol, Medvecký Martin, Trúchly Peter, Vargic Radoslav, Londák Juraj,  
Kadlic Radovan, Schumann Sebastian, Salazar Jordi, Silvestre Santiago, Levický  
Dušan, Polakovič Adam

Project website: <http://www.newtonproject.eu/>

This project has received funding from the European Union's Horizon 2020 Research and  
Innovation programme under Grant Agreement no. 688503.

# Index

<b>1. Introduction .....</b>	<b>4</b>
<b>2. NGN architecture evolution towards Future network architectures .....</b>	<b>5</b>
2.1 Multimedia services in NGN and Future networks environment.....	5
2.2 NGN concepts and architectures .....	6
<b>3. Advanced networks (Future networks) high level architecture.....</b>	<b>8</b>
<b>4. RM OSI and TCP/IP models.....</b>	<b>10</b>
4.1 Basic characteristic of network models.....	10
4.2 TCP/IP protocols .....	11
<b>5. CDN network architectures.....</b>	<b>23</b>
5.1 Nowadays CDNs in World.....	24
5.2 Content Flow .....	24
5.3 Control Layer .....	24
5.4 Distribution Layer .....	25
5.5 Federated CDNs .....	27
<b>6. Cloud computing .....</b>	<b>29</b>
6.1 Introduction. What is cloud computing? .....	29
6.2 History .....	30
6.3 Characteristics of cloud computing.....	32
6.4 Cloud computing components and architecture .....	33
6.5 Service models .....	34
6.6 Deployment models.....	39
6.7 Uses and applications .....	41
6.8 Benefits and disadvantages of cloud computing .....	42
6.9 Cloud security (potential privacy risks) .....	44
6.10 Conclusions .....	45
<b>7. Network virtualisation - SDN&amp;NFV .....</b>	<b>46</b>
7.1 SDN.....	46
7.2 NFV .....	52
<b>8. New generation of multimedia services/applications .....</b>	<b>59</b>
8.1 Introduction .....	59
8.2 Internet multimedia services and applications .....	59
8.3 Hybrid broadband broadcast TV services .....	61
8.4 eServices and mServices .....	63
8.5 Internet of things applications and services .....	65

8.6	NGN services .....	67
8.7	WebRTC.....	69
<b>9.</b>	<b>Information and network security.....</b>	<b>71</b>
9.1	Terminology in information security .....	71
9.2	Security mechanisms.....	73
9.3	Security attacks.....	74
9.4	Intruders.....	75
9.5	Malicious software .....	76
9.6	Network security models.....	77
9.7	Firewalls .....	79
9.8	Firewall categorization .....	79
9.9	Intrusion detection and prevention systems .....	82
<b>10.</b>	<b>Content adaptation.....</b>	<b>84</b>
10.1	Video quality and adaptation levels .....	84
10.2	Content adaptation in static image and audio transmission .....	85
10.3	Content adaptation in video transmission .....	87
<b>11.</b>	<b>5G Network architectures, services and applications.....</b>	<b>91</b>
11.1	Introduction .....	91
11.2	5G system architecture .....	92
11.3	5G services .....	98
11.4	5G applications.....	99

# 1. Introduction

By mutually cooperating different types of networks and their gradual integration into one universal broadband multimedia network - *Future Network (FN)*, the conditions for the transfer of all types of media and the provision of a wide range of multimedia services and applications are gradually being created. The **NGN** (*Next Generation Networks*) concept has evolved over several years, mainly within the working groups of the ITU and ETSI standardization institutions.

In the area of information and communication technologies, a continuous process of evolution is taking place, including the evolution of NGN technologies towards future networks (Future Network). There are several aspects that affect the current IMS-based NGN architecture in order to extend it to other features to support the implementation and delivery of next-generation multimedia services and applications.

This course offers selected topics covering various technologies that can be integrated within the new (future) generation networks (FN), i.e. new information and communication technologies as well as new services and applications.

The concept and structure of the course (study material) is designed so that it can be used both for the on-going form of education, but also as an online course on the basis of e-learning. The course uses new educational methodologies and new *information and communication technologies (ICT)* such as:

- "Self-directed learning" where learners themselves manage the learning process, with access to various learning resources, mainly via the Internet (educational portals, e-libraries, lecturer / tutors, discussion groups, social networks, experiments in virtual / fab laboratories, etc.)
- "Gamification" (learning games: 2D, 3D, with the possibility to use 3D VR applications - virtual reality, or 3D AR - augmented reality)
- Implementation of practical laboratory experiments in a virtual laboratory based on remote access.

## 2. NGN architecture evolution towards Future network architectures

The rapid development of Internet services and content delivery services across heterogeneous networks changes requirements from various aspects such as mobility, virtualization and resource sharing, security (network and information), simplification of architecture and flexibility in network management models, All this puts new demands on the evolution of NGN architectures toward FN (Future Networks). New topic in the area of information and communication technologies is further evolution of NGN (Next Generation Networks) technologies towards Future Networks.

### 2.1 Multimedia services in NGN and Future networks environment

New ICT-based concepts and architectures, based on converged ICT and NGN, offer operators new opportunities for implementing and delivering a broad spectrum of multimedia services and applications [1].

Currently HBB TV services are provided based on different technologies and platforms. Proposed evolution scenarios are focused on increasing of integration capabilities and compatibility. Proposed concepts and HBB TV systems migration scenarios are in relation with evolution processes of NGN towards FN.

Working groups of international standardization organizations ITU-T and ETSI according to ICT evolution are proposing new concepts and scenarios of multimedia services and applications.

It is appropriate to look into the evolution and to outline of NGN future trends and the open issues to be solved as well. New concepts and architectures of new generation of ICT based on converged ICT and NGN offer to operators new opportunities to implement and provide wide spectrum of multimedia services and applications.

Therefore, operators can leave a vertical structure of architecture where each type of service has predefined access, transport, management, and application infrastructure for the service, and moves to a horizontally oriented architecture, independent of the service provided (Figure 2.1).

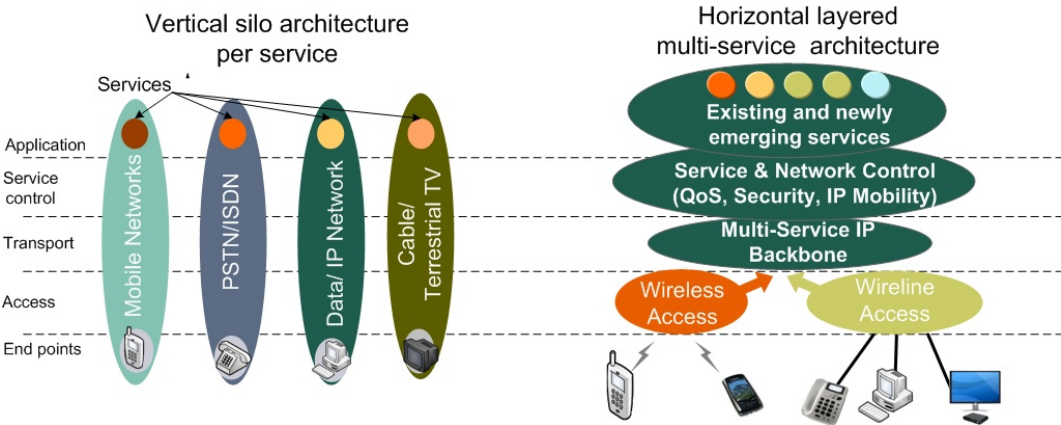


Figure 2.1 - From vertical silos to horizontal NGN architecture [2]

## 2.2 NGN concepts and architectures

The main principles of the NGNs (Next Generation Networks) were formed when the idea of NGN itself emerged. The next two definitions from ETSI and ITU-T describe NGN in substance.

ETSI describes NGN as a concept for the defining and establishing of the networks, allowing a formal distribution of functionalities into separate layers and planes by using open interfaces. The NGN concept provides new conditions for creation, implementation and effective management of innovative services.

ITU-T describes NGN as a network based on packet transfer, enabling to provide services, including telecommunication services, and is capable of using several broadband transmission technologies allowing guaranteeing QoS.

The functions related to services are at the same time independent of the basic transmission technologies. NGN provides unlimited user access to different service providers. It supports general mobility providing the users with consistency and availability of services.

That is what definitions say, but probably eventually NGN advantages are of bigger importance. Worth mentioning are some requirements for NGN it should conform to [3]:

- High-capacity packet transfer within the transmission infrastructure,
- Separation of managing functions from transmission features. Separation of service provisioning from the network,
- Support for a wide range of services and applications,
- Broadband capabilities, while complying with the requirements for **QoS** (*Quality of Services*),
- Various types of mobility (users, terminals, services),
- Various identification schemes and addressing,
- Converged services between fixed and mobile networks (as well as voice, data and video convergence),
- Conformance to the regulation requirements, such as emergency calls and security requirements,
- Cheaper and more effective technologies.

Within the NGN concepts the standardisation institutions are solving the following issues and problems:

- existing networks migration towards NGN,
- development in the field of access technologies,
- connection of other networks to IP networks,
- provision of services and development of new ones,

- interworking in the area of addressing,
- interworking of signalling systems,
- roaming a mobility.

There are many conceptual models and reference architectures for both the converged networks and VoIP architectures. Therefore, we have tried to find common features and to define a suitable conceptual model for NGN.

An objective of the conceptual model is to determine functional layers (covering similar functionalities), their entities, reference points (interfaces) and information flows between them. Such a model then can be mapped more easily into the physical reference architecture (and it is independent of the physical entities, i.e. components of the architecture).

In most analysed cases the NGN conceptual model layers are from the point of view of functionalities divided into independent parts as follows (Figure 2.2): access (some reference architectures do not include it directly into the NGN model or replace it by the adaptation one), transport (transmission, switching), control (call/sessions control) and application (services).

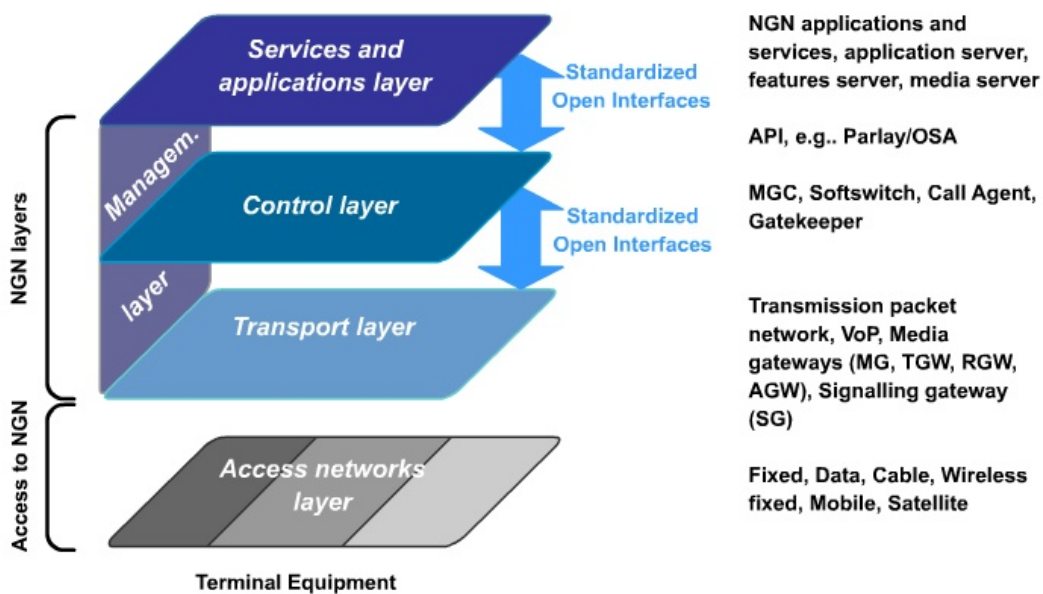


Figure 2.2 - NGN conceptual model and its functional layers [4]

### Conceptual model layers

*The access layer* provides the infrastructure, for example an access network between the end user and the transport network.

*The transport layer* ensures the transport between the individual nodes (points) of the network.

*The control layer* includes the control of services and network elements. This layer is responsible for set-up/establishing, control and cancelling of the multimedia session.

*The service/application layer* offers the basic service functions, which can be used to create more complex and sophisticated services and applications.

### 3. Advanced networks (Future networks) high level architecture

The architecture of the future network is illustrated in Fig. 3.1. It consists of several layers and planes.

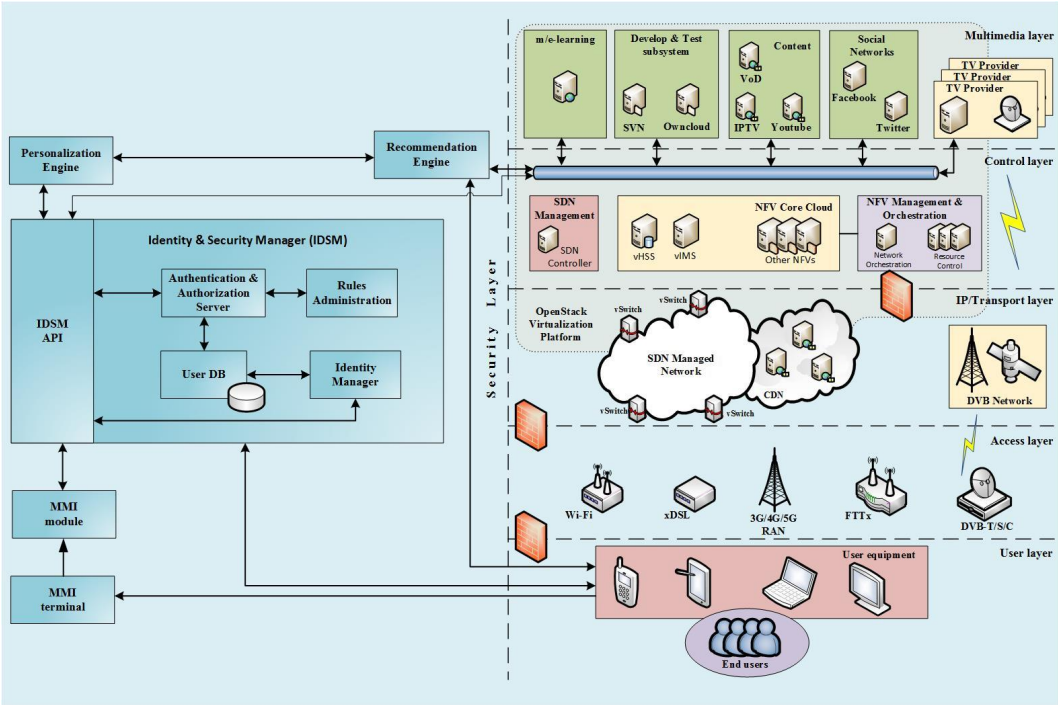


Figure 3.1 - Advanced Networks High Level Architecture

#### User Layer

The network will be used by end-users to communicate with each other or to obtain information and multimedia content provided through a multimedia layer. End-users can use different types of terminal equipment, such as phones, computers, game consoles, TV sets, with varying degrees of functionality (from ordinary telephone devices to smart phones with AV/VR support) or mobility (such as a PC or TV using a fixed home connection to portable laptops and mobile phones). In addition to people, the network will also be used for communication between different devices such as intelligent cars, various home furnishings (smart home), cities (smart city) and industry (Industrie 4.0). In the future, it is expected that a much larger number of devices will be connected to the network, and the amount of data transmitted by communication between devices will be more than the amount of data transmitted by people's communications.

#### Access Layer

The access layer provides connectivity for end-users and devices to the network. Similarly, as in the case of NGN, the access layer envisages the use of multiple progressive broadband technologies, both for solid optical or Metallic connection (FTTx, XDSL), as well as for wireless (WiFi, DVB-T/S) and mobile (3G/ 4G/5G) connection.



## **IP/Transport Layer**

The role of the transport layer is to ensure the transmission of data between terminal devices or between terminal devices and servers. The network will use optical, metal, microwave or satellite lines and various transport technologies (DWDM, OTH, MPLS, etc.). The uniting element will use the IP protocol as a single format for exchanging data.

The important features of the networks of the future are programmability and virtualization. The network will be formed on the principles of **SDN** (*Software defined Networking*). The SDN network uses network elements that have a separate data and control layer. The functionality of the control layer is centralized to the so-called SDN controller that controls the routing of packets in the network. The packet routing itself is performed in switches that are optimized for the performance of Forwarding packets and have limited functionality of the control part. For that reason, they can be simpler and cheaper. The above mentioned concept, described in detail in Chapter 7, enables flexible and efficient software networking and routing rules. In addition to the principles SDN can be part of architecture as well as other concepts such as the *Content Delivery Network* (**CDN**) used to distribute digital content (described in Chapter 5).

It is an effort to make the largest number of network functions (for routing, access control, etc.) not implemented through special, for the purpose of manufactured hardware devices but software, through applications crash the network Features in the cloud environment (see Chapter 6). This concept is yawned **NFV** (*Network Function Virtualization*) and is described in Chapter 7.

## **Control layer**

The control layer of the network provides management features of SDN network as well as the management of individual telecommunication services (e.g. IMS). The individual functional blocks providing control are understood as network functions and can therefore be implemented through virtualization by means of NFV (see Chapter 7).

An integral part of the control layer is also a function of management and orchestration of NFV infrastructure and services.

## **Multimedia layer**

The multimedia layer includes application and data servers providing functionality and data for individual services. Separating the service functionality into a single layer enables efficient and inexpensive provision of existing services and deployment of new ones. Similarly, as with the previous two layers and for the multimedia layer, the benefits of virtualization and cloudification can also be used to ensure the functionality of the multimedia layer.

## 4. RM OSI and TCP/IP models

### 4.1 Basic characteristic of network models

This section presents two network models first of which is currently used as a reference model (RM OSI) and second one is main network (protocol) model of current Internet (TCP/IP model).

#### 4.1.1 RM OSI Model

Designated ISO/IEC 7498-1 [5], the OSI model is a standard of the *International Organization for Standardization (ISO)*. It is a general-purpose paradigm for discussing or describing how computers communicate with one another over a network. Its seven-layered approach to data transmission divides the many operations up into specific related groups of actions at each layer (Fig. 4.1).

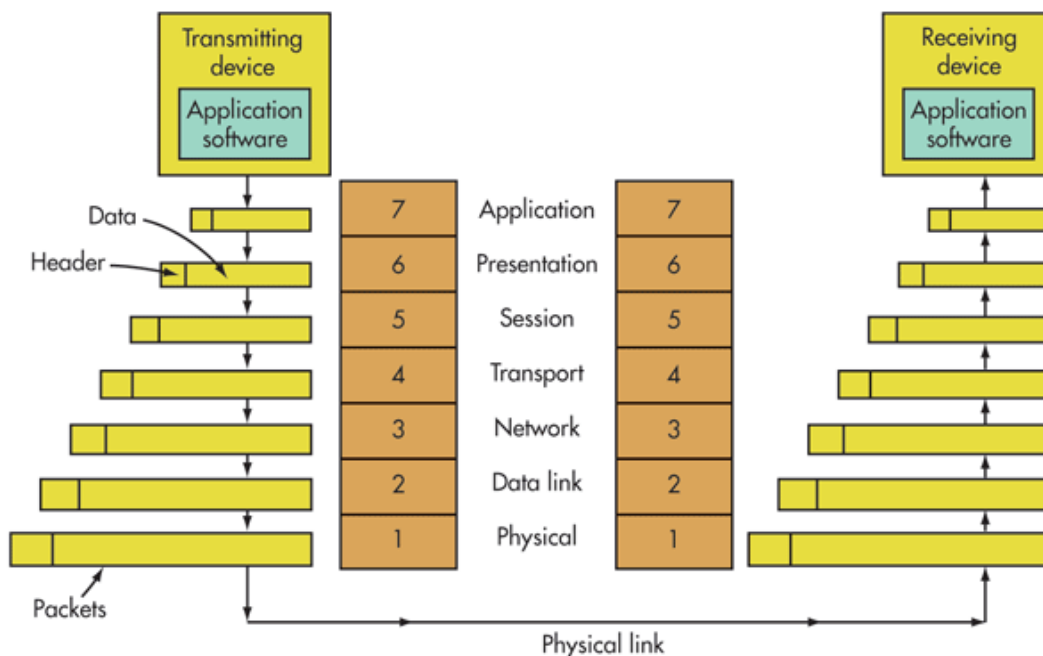


Figure 4.1 - Data flows in the OSI model

The principle of the OSI model is to divide the communication system into abstract layers. The original version of the model defined seven layers. At each level, the two communicating entities exchange the *protocol data units (PDUs)* through the protocols available in this layer. Each PDU contains a payload, called a *service data unit (SDU)*, along with the protocol header or footer.

Data flow between two OSI-compliant devices is as follows:

1. The transmitting data are combined on the top layer of the transmitting device (layer N) into the *protocol data unit (PDU)*.
2. The PDU goes into N-1 layer, where it is used as the *service data unit (SDU)*.

3. On layer N-1, the SDU merges with the N-1 PDU-forming header. It is then transferred to the N-2 layer.
4. The process continues until the lower level from which the data is transferred to the receiving device is reached.
5. On the receiving device, the data is transferred from the lowest to the top layer as the SDU series while being sequentially removed from the header or footer of each layer until the top layer where the last of the data is consumed.

#### 4.1.2 TCP/IP model

TCP/IP was developed during the 1960s as part of the Department of Defence's (DoD) Advanced Research Projects Agency (ARPA) effort to build a nationwide packet data network. It was first used in UNIX-based computers in universities and government installations. Today, it is the main protocol used in all Internet operations.

TCP/IP also is a layered protocol but does not use all of the OSI layers, though the layers are equivalent in operation and function (Fig. 4.2). The network access layer is equivalent to OSI layers 1 and 2. The Internet Protocol layer is comparable to layer 3 in the OSI model. The host-to-host layer is equivalent to OSI layer 4. These are the TCP and UDP (user datagram protocol) functions. Finally, the application layer is similar to OSI layers 5, 6, and 7 combined.

	OSI	TCP/IP
7	Application	Applications (FTP, SMTP, HTTP, etc.)
6	Presentation	
5	Session	
4	Transport	TCP (host-to-host)
3	Network	IP
2	Data link	Network access (usually Ethernet)
1	Physical	

Figure 4.2 - The seven layers of the OSI model somewhat correspond with the four layers that make up the TCP/IP protocol

## 4.2 TCP/IP protocols

This subchapter contains a brief description of protocols at the network, transport and application layer of the TCP/IP network model.

## 4.2.1 Network layer protocols

This section deals with following network layer protocols: IPv4, IPv6, ICMP, ICMPv6, IGMP and IPsec.

### IPv4

*Internet Protocol version 4 (IPv4)* is the fourth version of the *Internet Protocol (IP)* [6]. It is one of the core protocols of standards-based internetworking methods in the Internet. It still routes most Internet traffic today, despite the ongoing deployment of a successor protocol, IPv6. IPv4 is described in RFC 791.

IPv4 is a connectionless protocol for use on packet-switched networks.

It operates on a best effort delivery model in that it does not guarantee delivery, nor does it assure proper sequencing or avoidance of duplicate delivery. These aspects, including data integrity, are addressed by an upper layer transport protocol, such as the *Transmission Control Protocol (TCP)*.

IPv4 uses 32-bit addresses which limits the address space to 4294967296 ( $2^{32}$ ) addresses. IPv4 reserves special address blocks for private networks (~18 million addresses) and multicast addresses (~270 million addresses).

IPv4 addresses may be represented in any notation expressing a 32-bit integer value. They are most often written in the dot-decimal notation, which consists of four octets of the address expressed individually in decimal numbers and separated by periods.

The revised system of IPv4 addressing defines five classes. Classes A, B, and C had different bit lengths for network identification. The rest of the address is used to identify a host within a network, which meant that each network class had a different capacity for addressing hosts. Class D is defined for multicast addressing and Class E is reserved for future applications.

IPv4 address classes

Class	Address range	1st octet	Notes
A	0000000 - 01111111	1 - 127	16 777 214 hosts per subnet
B	1000000 - 10111111	128 - 191	65 534 hosts per subnet
C	1100000 - 11011111	192 - 223	254 hosts per subnet
D	1110000 - 11101111	224 - 239	reserved for Multicasting
E	1111000 - 11110111	240 - 255	reserved for experimental purposes

Of the approximately four billion addresses defined in IPv4, three ranges are reserved for use in private networks. Packets addresses in these ranges are not routable in the public Internet, because they are ignored by all public routers. Therefore, private hosts cannot directly communicate with public networks, but require network address translation at a routing gateway for this purpose.

#### Reserved address ranges for private addresses

Name	Address range	Number of addresses
24-bitový blok	10.0.0.0 - 10.255.255.255	16 777 216
20-bitový blok	172.16.0.0 - 172.31.255.255	1 048 576
16-bitový blok	192.168.0.0 - 192.168.255.255	65 536

A broadcast address is an address that allows information to be sent to all interfaces in a given subnet, rather than a specific machine.

Generally, the broadcast address is found by obtaining the bit complement of the subnet mask and performing a bitwise OR operation with the network identifier. In other words, the broadcast address is the last address in the address range of the subnet.

### IPv6

*Internet Protocol version 6 (IPv6)* is the most recent version of the *Internet Protocol (IP)* developed by the *Internet Engineering Task Force (IETF)* [7] to deal with the long-anticipated problem of IPv4 address exhaustion.

IPv6 uses a 128-bit address, theoretically allowing  $2^{128}$ , or approximately  $3.4 \times 10^{38}$  addresses. The actual number is slightly smaller, as multiple ranges are reserved for special use or completely excluded from use. The total number of possible IPv6 addresses is more than  $7.9 \times 10^{28}$  times as many as IPv4, which uses 32-bit addresses and provides approximately 4.3 billion addresses.

IPv6 addresses are represented as eight groups of four hexadecimal digits with the groups being separated by colons, for example 2001:0db8:0000:0042:0000:8a2e:0370:7334, but methods to abbreviate this full notation exist.

IPv6 is intended to replace IPv4 and these two protocols are not designed to be interoperable.

IPv6 provides other technical benefits in addition to a larger addressing space. In particular, it permits hierarchical address allocation methods that facilitate route aggregation across the Internet, and thus limit the expansion of routing tables. The use of multicast addressing is expanded and simplified and provides additional optimization for the delivery of services. Device mobility, security, and configuration aspects have been considered in the design of the protocol.

### IPv6 addressing

IPv6 addresses are assigned to interfaces, rather than to nodes, in recognition that a node can have more than one interface. Moreover, it can be assigned more than one IPv6 address to an interface.

IPv6 defines three address types:

- **unicast** - Identifies an interface of an individual node.
- **multicast** - Identifies a group of interfaces, usually on different nodes. Packets that are sent to the multicast address go to all members of the multicast group.

- **anycast** - Identifies a group of interfaces, usually on different nodes. Packets that are sent to the anycast address go to the anycast group member node that is physically closest to the sender.

An IPv6 address is 128 bits in length and consists of eight, 16-bit fields, with each field bounded by a colon. Each field must contain a hexadecimal number, in contrast to the dotted-decimal notation of IPv4 addresses.

The leftmost three fields (48 bits) contain the site prefix. The prefix describes the public topology that is usually allocated to your site by an ISP or *Regional Internet Registry (RIR)*.

The next field is the 16-bit subnet ID, which you (or another administrator) allocate for your site. The subnet ID describes the private topology, also known as the site topology, because it is internal to your site.

The rightmost four fields (64 bits) contain the interface ID, also referred to as a token. The interface ID is either automatically configured from the interface's MAC address or manually configured in EUI-64 format.

An example of an IPv6 address: 2001:0db8:3c4d:0015:0000:0000:1a2f:1a2b

## ICMP

**ICMP** (*Internet Control Message Protocol*) [8] is a supporting protocol in the Internet protocol suite. It is used by network devices, including routers, to send error messages and operational information indicating, for example, that a requested service is not available or that a host or router could not be reached. ICMP differs from transport protocols such as TCP and UDP in that it is not typically used to exchange data between systems, nor is it regularly employed by end-user network applications (with the exception of some diagnostic tools like ping and traceroute).

## ICMPv6

**ICMPv6** (*Internet Control Message Protocol version 6*) [9] is the implementation of the *Internet Control Message Protocol (ICMP)* for *Internet Protocol version 6 (IPv6)*. ICMPv6 is defined in RFC 4443. ICMPv6 is an integral part of IPv6 and performs error reporting and diagnostic functions (e.g., ping), and has a framework for extensions to implement future changes.

## IGMP

**IGMP** (*Internet Group Management Protocol*) [10] is a communications protocol used by hosts and adjacent routers on IPv4 networks to establish multicast group memberships. IGMP is an integral part of IP multicast.

IGMP can be used for one-to-many networking applications such as online streaming video and gaming, and allows more efficient use of resources when supporting these types of applications.

There are three versions of IGMP. These versions are backwards compatible, so IGMPv3 supporting routers can support clients with IGMPv1, IGMPv2, and IGMPv3.

- IGMPv1 uses a query response model. Queries are sent to 224.0.0.1. Membership updates are sent to the group multicast address.

- IGMPv2 speeds up the process of leaving the group and adjusts other timeouts. Messages leaving groups are sent to 224.0.0.2. A specific query for the group was introduced. Group queries are sent to the group multicast address.
- IGMPv3 introduces source-specific multicast functions. Membership updates are sent to 224.0.0.22

## IPsec

**IPsec** (*Internet Protocol Security*) is a network protocol suite that authenticates and encrypts the packets of data sent over a network. IPsec includes protocols for establishing mutual authentication between agents at the beginning of the session and negotiation of cryptographic keys to use during the session.

IPsec can protect data flows between a pair of hosts (host-to-host), between a pair of security gateways (network-to-network), or between a security gateway and a host (network-to-host). IPsec uses cryptographic security services to protect communications over IP networks.

IPsec is an end-to-end security scheme operating in the Internet layer of the Internet protocol suite, while some other Internet security systems in widespread use, such as *Transport Layer Security* (TLS) and *Secure Shell* (SSH), operate in the upper layers at the transport layer (TLS) and the application layer (SSH).

Hence, only IPsec protects all application traffic over an IP network. IPsec can automatically secure applications at the IP layer.

### 4.2.2 Transport layer protocols

This section deals with following transport layer protocols: TCP, UDP, SCTP and RSVP.

#### TCP

**TCP** (*Transmission Control Protocol*) [11] is one of the main protocols of the Internet protocol suite. TCP is a connection-oriented protocol, which means that it creates and maintains a connection until the programs at both ends of the communication do not complete the message exchange. TCP is responsible for dividing application data into segments so that they can be delivered via the network, managing stream management, and processing retransmission of lost or defective packets as well as confirming all received packets.

Because of its focus to precise over timed delivery orientation, it can cause relatively long delays (in seconds) caused by waiting for unattended messages or retransmissions of lost messages. Therefore, it is not suitable for real-time applications such as Voice over IP.

A three-way handshake is used to create a TCP connection. Before the client attempts to connect to the server, they must first connect and listen to the specified port to open it for connection - this is called passive opening. After creating a passive opening, the client can initiate active opening.

#### UDP

**UDP** (*User Datagram Protocol*) [12] uses a simple connectionless transmission model with a minimum of protocol mechanism. It uses checksums to check the integrity of data and port numbers to address the various functions at the source and target station.

No connection-oriented mechanisms have been implemented in UDP protocol. Therefore, it does not provide any check of the accuracy of the delivered data and there is no guarantee of delivery, order or duplicate protection.

With this simplicity, UDP is a convenient protocol for streaming media and real-time data streaming.

## **SCTP**

**SCTP** (*Stream Control Transmission Protocol*) [13] is a transport-layer protocol, serving in a similar role to the popular protocols TCP and UDP. It is standardized by IETF in RFC 4960. SCTP provides some of the same service features of both UDP and TCP: it is message-oriented like UDP and ensures reliable, in-sequence transport of messages with congestion control like TCP; it differs from these in providing multi-homing and redundant paths to increase resilience and reliability.

In the absence of native SCTP support in operating systems it is possible to tunnel SCTP over UDP, as well as mapping TCP API calls to SCTP ones.

## **RSVP**

**RSVP** (*Resource Reservation Protocol*) [14] is a transport layer protocol designed to reserve resources across a network for an integrated services Internet. RSVP operates over an IPv4 or IPv6 Internet Layer and provides receiver-initiated setup of resource reservations for multicast or unicast data flows with scaling and robustness. It does not transport application data but is similar to a control protocol, like ICMP or IGMP. RSVP is described in RFC 2205.

RSVP can be used by either hosts or routers to request or deliver specific levels of quality of service (QoS) for application data streams or flows. RSVP defines how applications place reservations and how they can relinquish the reserved resources once the need for them has ended. RSVP operation will generally result in resources being reserved in each node along a path.

RSVP is not a routing protocol and was designed to interoperate with current and future routing protocols. RSVP by itself is rarely deployed in telecommunications networks today but the traffic engineering extension of RSVP, or RSVP-TE, is becoming more widely accepted nowadays in many QoS-oriented networks.

### 4.2.3 Application layer protocols

This section deals with following application layer protocols: DHCP, DNS, TLS/SSL, Telnet, FTP, HTTP, POP, IMAP, SMTP, XMPP, SOAP, RIP, SIP, OSPF, IS-IS, RTP, RTCP, RTSP, SNMP and NETCONF.

## **DHCP**

**DHCP** (*Dynamic Host Configuration Protocol*) [15] is a network protocol used to manage TCP/IP settings (IP addresses, default gateway, domain name, the name servers, and time servers) for devices on a network. The DHCP uses a client-server model. The DHCP server manages a pool of IP addresses and information about client configuration. When a computer or other device connects to a network, the DHCP client sends a broadcast query requesting the necessary information. The client broadcasts messages on the network subnet using the destination address 255.255.255.255 or the specific subnet broadcast address. Any DHCP



server on the network may service the request. When a DHCP server receives an IP address lease request message from a client, it reserves an IP address for the client and makes a lease offer by sending a message containing the client's MAC address, the offered IP address, the subnet mask, the lease duration, and the IP address of the DHCP server making the offer to the client. A client can receive DHCP offers from multiple servers, but it will accept only one of them. In response to the DHCP server offer, the client broadcast to the server requesting the offered address. When the DHCP server receives the request message from the client, it sends an acknowledgement to the client. The DHCP uses a connectionless service model, using the UDP port 67 for server and UDP port 68 for client. Depending on implementation, the DHCP server may have three methods of allocating IP addresses:

- *Dynamic allocation* - DHCP server assign IP addresses for specific time period and then reallocate IP addresses that are not renewed.
- *Automatic allocation* - the DHCP server keeps a table of past IP address assignments, and preferentially assigns to a client the same IP address that the client previously had.
- *Manual (static) allocation* - the DHCP server assigns a private IP address based on a predefined mapping by the administrator.

## **DNS**

**DNS** (*Domain Name System*) [16] is a hierarchical decentralized naming system for the translation of human-friendly computer hostnames into numeric IP addresses. The DSN provides a worldwide, distributed directory service for the Internet.

The DNS protocol defines a detailed specification of the data structures and data communication exchanges used in the DNS. The DNS protocol uses two types of DNS messages, queries and replies. DNS uses the UDP (port number 53) to serve requests. DNS queries consist of a single UDP request from the client followed by a single UDP reply from the server. The TCP is used in special cases, e.g. when the response data size exceeds 512 bytes.

## **TLS/SSL**

**TLS** (*Transport Layer Security*) [17] and its predecessor, **SSL** (*Secure Sockets Layer*) [18], both frequently referred to as "SSL", are cryptographic protocols that provide communications security over a computer network. Several versions of the protocols are used in applications such as web browsing, email, Internet faxing, instant messaging, and voice-over-IP (VoIP). Websites are able to use TLS to secure all communications between their servers and web browsers. The TLS protocol comprises two layers: the TLS record protocol and the TLS handshake protocol. TLS supports many different methods for exchanging keys, encrypting data, and authenticating message integrity.

## **Telnet**

**Telnet** [19] is a client-server oriented communication protocol used for bidirectional interactive text-oriented communication over virtual terminal connection.

Telnet was used to provide access to a command-line interface (CLI) on a remote host, but due to security issues (Telnet, by default, does not encrypt any data sent over the connection, including passwords) when used over an open network such as the Internet, it is replaced by SSH.

Telnet uses a TCP port number 23.

## **FTP**

**FTP** (*File Transfer Protocol*) [20] is a standard network protocol used for the transfer of computer files between a client and server on a computer network. FTP is built on a client-server model architecture and uses separate control and data connections between the client and the server.

FTP users may authenticate themselves with a clear-text sign-in protocol, normally in the form of a username and password, but can connect anonymously if the server is configured to allow it. For secure transmission that protects the username and password, and encrypts the content, FTP is often secured with SSL/TLS (FTPS) [21].

## **HTTP**

**HTTP** (*Hypertext Transfer Protocol*) [22] is an application protocol for distributed, collaborative, and hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web.

Hypertext is structured text that uses logical links (hyperlinks) between nodes containing text and HTTP is the protocol to exchange or transfer hypertext.

HTTP resources are identified and located on the network by *Uniform Resource Locators (URL)* [23], using the *Uniform Resource Identifiers (URI)* schemes which defines the domain address of the server, the source location on the server, and the protocol that can be used to access the source, e.g. <https://tools.ietf.org/html/rfc2616>.

## **POP**

**POP** (*Post Office Protocol*) [24] is an application-layer Internet standard protocol used by local e-mail clients to retrieve e-mail from a remote server over a TCP/IP connection. POP has been developed through several versions, with version 3 (POP3) being the last standard in common use before largely being made obsolete by the more advanced IMAP as well as webmail.

## **IMAP**

**IMAP** (*Internet Message Access Protocol*) [25] is an Internet standard protocol used by e-mail clients to retrieve e-mail messages from a mail server over a TCP/IP connection.

IMAP was designed with the goal of permitting complete management of an email box by multiple email clients, therefore clients generally leave messages on the server until the user explicitly deletes them. An IMAP server typically listens on port number 143. IMAP over SSL (IMAPS) is assigned the port number 993.

Virtually all modern e-mail clients and servers support IMAP. IMAP and the earlier **POP3** (*Post Office Protocol*) are the two most prevalent standard protocols for email retrieval.

## **SMTP**

**SMTP** (*Simple Mail Transfer Protocol*) [26] is an Internet standard for electronic mail (email) communication between electronic mail servers. Although electronic mail servers and other mail transfer agents use SMTP to send and receive mail messages, user-level client mail

applications typically use SMTP only for sending messages to a mail server for relaying. For retrieving messages, client applications usually use either IMAP or POP3.

## **XMPP**

**XMPP** (*Extensible Messaging and Presence Protocol*) [27] is a communications protocol for message-oriented middleware based on **XML** (*Extensible Markup Language*) [28]. It enables the near-real-time exchange of structured yet extensible data between any two or more network entities.

Designed to be extensible, the protocol has been used also for publish-subscribe systems, signalling for VoIP, video, file transfer, gaming, the *Internet of Things (IoT)* applications such as the smart grid, and social networking services.

## **SOAP**

**SOAP** (*Simple Object Access Protocol*) [29] is a protocol specification for exchanging structured information in the implementation of web services in computer networks. Its purpose is to induce extensibility, neutrality and independence. It uses XML Information Set for its message format, and relies on application layer protocols, most often *Hypertext Transfer Protocol (HTTP)* or *Simple Mail Transfer Protocol (SMTP)*, for message negotiation and transmission.

SOAP allows processes running on disparate operating systems (such as Windows and Linux) to communicate using *Extensible Markup Language (XML)*. Since Web protocols like HTTP are installed and running on all operating systems, SOAP allows clients to invoke web services and receive responses independent of language and platforms.

## **SIP**

**SIP** (*Session Initiation Protocol*) [30] is a communications protocol for signalling, for the purpose of controlling multimedia communication sessions. The most common applications of SIP are in Internet telephony for voice and video calls, private IP telephone systems, as well as instant messaging over Internet Protocol (IP) networks.

The protocol defines the messages that are sent between endpoints, which govern establishment, termination and other essential elements of a call. SIP can be used for creating, modifying and terminating sessions consisting of one or several media streams.

SIP is designed to be independent (although not agnostic) of the underlying transport layer, and can be used with UDP, TCP, and SCTP and can be secured using TLS. It is a text-based protocol, incorporating many elements of the HTTP and SMTP.

By itself, SIP only provides signalling; it is used in conjunction with other protocols that specify the media format and protocol to be used to subsequently communicate the media. Although SIP can carry arbitrary data, SIP is typically used to carry a *Session Description Protocol (SDP)* message specifying the codec and the use of either the *Real-time Transport Protocol (RTP)* or *Secure Real-time Transport Protocol (SRTP)* for (media) communication.

## **RIP**

**RIP** (*Routing Information Protocol*) [31] is dynamic distance-vector routing protocol which employ the hop count as a routing metric. RIP prevents routing loops by implementing a limit on the number of hops allowed in a path from source to destination. The maximum number of

hops allowed for RIP is 15, which limits the size of networks that RIP can support. A hop count of 16 is considered an infinite distance and the route is considered unreachable.

In most networking environments, RIP is not the preferred choice for routing as its time to converge and scalability are poor compared to EIGRP, OSPF, or IS-IS.

However, it is easy to configure, because RIP does not require any parameters unlike other protocols.

## **OSPF**

**OSPF** (*Open Shortest Path First*) [32] is a routing protocol for *Internet Protocol (IP)* networks. It uses a link state routing algorithm and falls into the group of *interior gateway protocols (IGP)*, operating within a single *autonomous system (AS)*.

OSPF can detect changes in network topology (e.g., link break) and reconfigure network topology into a new structure within seconds. OSPF calculates the shortest path for each direction using the Dijkstra algorithm. OSPF is the most widely used IGP routing protocol in large enterprise networks.

Currently used OSPF versions are OSPFv2 for IPv4 networks and OSPFv3 [33] for IPv6 networks.

## **IS-IS**

**IS-IS** (*Intermediate System to Intermediate System*) [34] is a link-state dynamic routing protocol used in large service provider networks. IS-IS is an interior gateway protocol, designed for use within an administrative domain or network. This is in contrast to exterior gateway protocols, primarily *Border Gateway Protocol (BGP)*, which is used for routing between autonomous systems

Like OSPF, it routes according to the link state and uses the Dijkstra algorithm to calculate the best path in the network. Unlike the OSPF that was proposed for IP routing on the 3rd OSI Layer, the IS-IS protocol was designed as an ISO standard for OSI Layer 2 routing and does not use IP to transmit routing information. IS-IS is neutral from the point of network addresses, making it easy to deploy for routing in IPv6 networks. For this reason, IS-IS is considered as de facto standard in the core networks of major Internet providers.

## **BGP**

**BGP** (*Border Gateway Protocol*) [35] is a standardized exterior gateway protocol designed to exchange routing and reachability information among *autonomous systems (AS)* on the Internet. The Border Gateway Protocol makes routing decisions based on paths, network policies, or rule-sets configured by a network administrator and is involved in making core routing decisions.

BGP may be used for routing within an autonomous system. In this application it is referred to as Interior Border Gateway Protocol, Internal BGP, or **iBGP** (*Interior Border Gateway Protocol*). In contrast, the Internet application of the protocol may be referred to as Exterior Border Gateway Protocol, External BGP, or **eBGP** (*Exterior Border Gateway Protocol*). The current version is BGP4.

## **RTP**

**RTP** (*Real-time Transport Protocol*) [36] is a network protocol for delivering audio and video over IP networks. RTP is used extensively in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications, television services and web-based push-to-talk features.

RTP typically runs over *User Datagram Protocol* (**UDP**). RTP is used in conjunction with the *RTP Control Protocol* (**RTCP**). While RTP carries the media streams (e.g. audio and video), RTCP is used to monitor transmission statistics and *quality of service* (**QoS**) and aids synchronization of multiple streams. RTP is one of the technical foundations of Voice over IP and in this context is often used in conjunction with a signalling protocol such as the *Session Initiation Protocol* (**SIP**) which establishes connections across the network.

## **RTCP**

**RTCP** (*RTP Control Protocol*) [37] is a sister protocol of the *Real-time Transport Protocol* (**RTP**). Its basic functionality and packet structure is defined in RFC 3550. RTCP provides out-of-band statistics and control information for an RTP session. It partners with RTP in the delivery and packaging of multimedia data, but does not transport any media data itself. The primary function of RTCP is to provide feedback on the *quality of service* (**QoS**) in media distribution by periodically sending statistics information to participants in a streaming multimedia session.

RTCP transports statistics for a media connection and information such as transmitted octet and packet counts, packet loss, packet delay variation, and round-trip delay time. An application may use this information to control quality of service parameters, perhaps by limiting flow, or using a different codec.

## **RTSP**

**RTSP** (*Real Time Streaming Protocol*) [38] is a network control protocol designed for use in entertainment and communications systems to control streaming media servers. The protocol is used for establishing and controlling media sessions between end points. Clients of media servers issue VCR-style commands, such as play, record and pause, to facilitate real-time control of the media streaming from the server to a client (Video on Demand) or from a client to the server (Voice Recording).

The transmission of streaming data itself is not a task of RTSP. Most RTSP servers use the *Real-time Transport Protocol* (**RTP**) in conjunction with *Real-time Control Protocol* (**RTCP**) for media stream delivery. However, some vendors implement proprietary transport protocols.

## **SNMP**

**SNMP** (*Simple Network Management Protocol*) [39] is an Internet-standard protocol widely used in network management for network monitoring. SNMP exposes management data in the form of variables on the managed systems organized in a *management information base* (**MIB**) [40] which describe the system status and configuration. These variables can then be remotely queried (and, in some circumstances, manipulated) by managing applications.

Three significant versions of SNMP have been developed and deployed. SNMPv1 is the original version of the protocol. More recent versions, SNMPv2c [41] and SNMPv3 [42], bring feature improvements in performance, flexibility and security.

## NETCONF

**NETCONF** (*Network Configuration Protocol*) [43] is a network management protocol developed and standardized by the IETF. NETCONF provides mechanisms to install, manipulate, and delete the configuration of network devices. Its operations are realized on top of a simple *Remote Procedure Call (RPC)* layer. The NETCONF protocol uses an *Extensible Markup Language (XML)* based data encoding for the configuration data as well as the protocol messages. The protocol messages are exchanged on top of a secure transport protocol.

## 5. CDN network architectures

*Content Delivery Network (CDN)* is a network, which consists of huge number of distribution points (called nodes). This network is running over the Internet network and can defectively deliver huge amount of content to huge amount of recipients.

The most popular usage of these networks is distribution of software updates.

CDN is owned by a CDN provider, which is responsible for content distribution, which is ingested by a CDN customer.

The advantage of CDN is in resource sharing, because if every software developer wants to distribute updates over his customers in Internet, he must build system of servers, maintain and operate its. Then there will be numbers of parallel systems to the same point. This overlap and low utilization of nodes will increase the price/usability ratio. Distribution from one central point will use huge amount of capacity for transferring the same thing many times.

CDN network is organized in two layers:

- control layer - mostly central node, which is used for content ingestion and distribution over the distribution layer. It is also responsible for Authentication and Authorization functionality if implemented.
- distribution layer - forest of nodes, distributed over many locations.

CDN distribution is split in two methods:

1. on-line - is the same as multicast, all recipients get the same data at same time.

Good example is sport match, when every customer wants to see the match with minimal delay (acceptable value is tens of seconds).

2. off-line - information which generation has started and ended in the past. Off-line distribution can be split in two types of delivery methods: download, streaming.

Example can be well known video stream (YouTube) or distribution of updates.

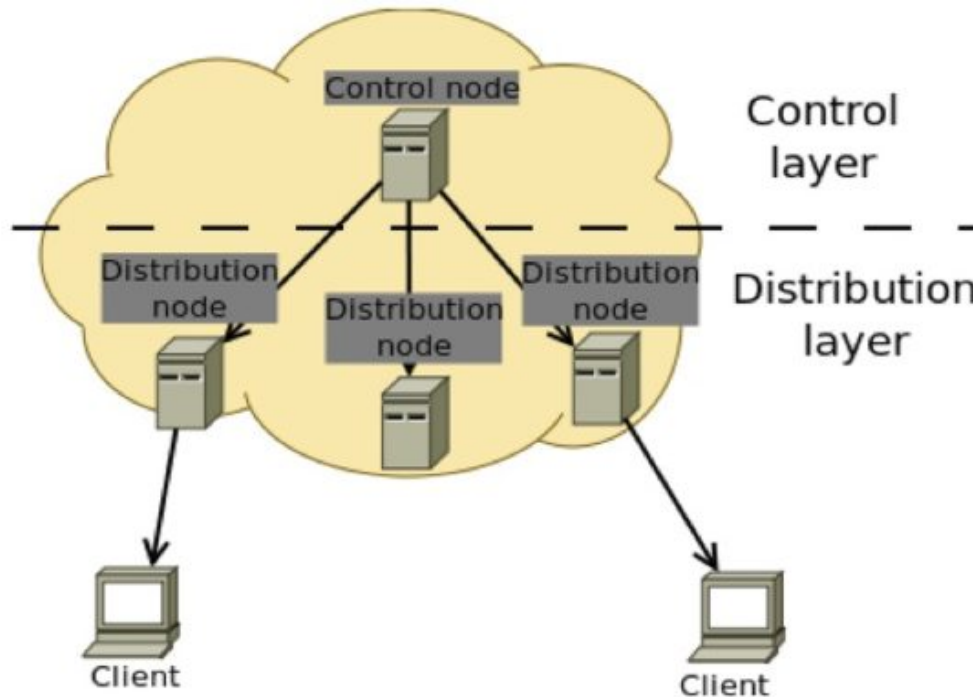


Figure 5.1 - CDN Architecture

## 5.1 Nowadays CDNs in World

CDNs are now currently implemented as a separate commercial product. To provide unique functionality and effective operation, CDN operators have implemented their own algorithms and transfer protocols. There also exists commercial software for CDN creation at own premises of Telecommunication operators and other companies interested in own CDN implementation, e.g. EdgeCast [44].

Open-source projects are based on activities of single projects, which have developed their own CDN networks based on open source software. Every volunteer can join this network with his own server or cluster. This new server will become a new node of open CDN and will be controlled by this open CDN network management.

## 5.2 Content Flow

Nowadays, the Internet content based service are being split in main parts: *Content*, *Service*, *Network (Transport)*, *Consumer*. All these entities have clear relations among them, which defines clear content flow from a content originator (author) to a content consumer. These parts are provided and represented by adequate entities. Content is originated from the author. Then the author passes content with right to a content provider. The content provider will then provide content to service operators. Service operators are offering service consisting of content to the end customer via network. Network is provided by a network operator. Not all the entities are necessarily separated from each other. Mostly, the service provider and network operator are the same entity or the content provider also operates the service.

## 5.3 Control Layer

This layer is responsible for a management of content distribution, content ingestion, reporting, logging, request routing function.



Every CDN have some capabilities for content delivery methods, content adaptation and content security. All these information are very important for content transfer from a source to the consumer.

- Delivery methods - by delivery methods, we understood different methods for content transfer from distribution node to end consumer.
- Content adaptation - is set of methods, which can be used to change original information to form, which is more useful for transfer over real-time channel.
- Content security - is set of rules and mechanisms for protection of content including access authorization, digital rights management, copy protection, watermarking, ...

It is very important to map, which area can be served from which distribution node or network. In IP network this is done by a routing table, where every item - subnet is described by a network address and network mask. In CDN networks the address and size is defined in one parameter called a footprint. The footprint can be potentially any information about coverage like geographical area [45], Internet provider [46], etc. There exists a footprint table which contains list of all footprints with their distribution nodes.

## **5.4 Distribution Layer**

This layer is responsible for content delivery to the consumer. It consists of huge number of delivery nodes, which can be organized also in clusters. We can call these nodes or clusters as distribution points.

Every distribution point contains big storage capacity on local or external disk arrays. This capacity is used for the content storage.

### **5.4.1 File based content**

File based content is sequence of bytes which have started and ended in the past.

File based content can be downloaded without QoS, but almost always requires zero tolerance to change or modification at any level of content.

This type of content is majority of all content delivered over the Internet and also over the CDNs. For this purpose the ideal transfer protocol is HTTP, which is widely used in current Internet.

### **5.4.2 Stream based content**

A stream is sequence of bytes which started and ended in the past.

The stream based content contains useful information per parts, so it can be transferred continuously and during the transfer it can be also consumed. Streaming content is mostly sensitive to QoS of transfer channel, especially to delay and jitter.

Typical content of this type is a media stream. In contrast to data transfers in Internet the video and audio content is becoming more and more popular. CDN like part of Internet is helping to fulfil these requirements by spreading the content and the load over more location and distribution points.

Media consumers are sensitive to audio information, because every country is using different language. To save resources in the network it is effective to transfer only that audio stream which contains required language. To do this a distribution point can contain content with video stream and multiple audio streams. Only one audio stream (elementary stream) is transferred to consumer in a session according consumer set up. Every elementary stream can be transferred over different paths from source to destination.

The same concept can be also used for multiple angle content, where single content is recorded from more video cameras from different angles (views). Again only video is transferred from distribution point to consumer.

### 5.4.3 Live stream based content

This stream is a sequence of bytes which have started in the past and is not finished until now and continues.

Because this stream is live, it cannot be cached for further streaming and all the efficiency in content distribution is in single transfer over the same parts of network.





Figure 5.2 - Live stream distribution without CDN and with CDN distribution

More clear explanation would be on simple example: imagine that two consumers in Europe want to see live video show from China Olympic Games. In a normal unicast network, which Internet is, the stream will be duplicated at China at a streaming platform and transferred to consumers in two separate streams (Figure 5.2).

By redirection of consumers to CDN, CDN can join the video stream from China Streaming server, transfer it to its location in Europe (Figure 5.2). The consumers then connect to CDN and stream is duplicated in CDN distribution point and transferred to the consumers in two separate streams, but only inside the European network. This concept can save Internet connectivity from Europe to China. In this case bandwidth is 1/2 of previous concept. If there are three consumers of same content saving is 1/3 of previous concept, etc.

## 5.5 Federated CDNs

It is very hard to deploy global international network, which is split over the whole globe. CDN benefits from huge number of delivered content (bytes). More bytes delivered by certain platform are increasing utilization of network; therefore the price of one transferred byte is lower. Lower price per byte brings more customers, which brings more number of content which is delivered to the customers. More content will consume more bytes for delivery. As we can see the pricing process is rounded in one loop without small external impact. This loop is “responsible” for creation of big CDN players, which are controlling the market.

Telecommunication operators are losing market share day by day and international global service providers (Akamai, Globix, Limelight ...) are overtaking the service market. Telecommunication operators are pushed to role of a network provider “cable owner”. This situation is unwanted for Telecommunication operator and therefore they are investing afford to value add services like CDNs.

Under CDN federation, there is huge number of initiatives driven by CDN owners and mostly by telecommunication operators [47], [48]. Federation is way of CDN interconnection, where

content can be copied from one CDN to another. Then, the end user can be served by both CDNs by content originated network or by federated network.

Few federated networks act and look like one robust CDN from user perspective and from content provider perspective.

## 6. Cloud computing

### 6.1 Introduction. What is cloud computing?

According to NIST, cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [49]. A complementary definition is provided by the RAD Lab at the University of Berkeley [50]. The authors consider that cloud computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centres that provide those services. Basically, cloud computing is a distributed computing paradigm that focuses on providing a wide range of users with distributed access to scalable, virtualized hardware and/or software infrastructure over Internet.

Cloud computing refers to the delivery of computing resources over the Internet. It consists of a collection of technologies that ensure, typically in the form of a service offered by a provider to the customer, a consistency storage or high-performance data processing capabilities through the use of hardware/software distributed and virtualized network. The cloud provider can both own and house the hardware and software necessary to run the home or business applications.

Cloud computing is a complex and rapidly evolving concept. Cloud computing may be regarded as a distributed system that offers computing services via a computer communication network, usually the Internet. Resources in the cloud are transparent to the users, and the users do not need to know their exact location. They can be shared among a large number of users, who should be able to access applications and data from anywhere at any time.

A simple example of cloud computing is webmail. The webmail provider maintains the server space and provides access; the webmail user just plugs a web address into a browser and submits user information to access an account. The software and storage for her account does not exist on her computer - it's on the service's computer cloud.

The main objective of cloud computing is to make better use of these distributed resources and solve large-scale computation problems. The word "cloud" is a metaphor for describing the Web as a space where computing has been preinstalled and exists as a service [49]. Operating systems, applications, storage, data, and processing capacity all exist on the Web, ready to be shared among users. The use of cloud computing entails that a user can benefit from data processing resources and storage that a company offers her as a service over the Internet, instead of using only the hardware and software of her computer or servers located within the corporate network. Generally, these services are offered in a completely transparent manner; the platforms hide the complexity and details of the underlying infrastructure from users and applications. In the cloud computing model, computing power, software, storage services, and platforms are delivered on demand to external customers over the Internet. Potentially, all kind of applications from word processing software to customized computer programs could work on a cloud computing system. The access that this technology provides to resources and services can be scaled up or down to meet demand. Cloud computing providers typically charge customers on a pay-per-use model.

While the benefits of cloud computing are many, its disadvantages are just as numerous. If used properly, cloud computing is a technology with great opportunity for businesses of all sizes.

The major advantages of cloud computing are on-demand self-service, ubiquitous network access, location-independent resource pooling, and transference of risk. Additional advantages include lower running costs, ease of utilization, quality of service, and reliability, no large-scale infrastructure investment, better agility and scalability and better peak demand management. For example, cloud computing can focus the power of thousands of computers on one problem, enabling researchers to do their work faster than ever.

Nowadays, the greatest challenges with cloud computing are privacy and security. Other disadvantages are the lack or limited control, the implicit dependency on the provider also known as "vendor lock-in". It is difficult to migrate from a provider to another once a user has rolled with it.

Vivek Kundra, Federal CIO and formerly technology chief of District of Columbia said: "The cloud will do for government what the Internet did in the '90s," [51].

## 6.2 History

Information Technology has always been considered a major pain point of enterprise organizations, from the perspectives of both cost and management. However, the information technology industry has experienced a dramatic shift in the past decade - factors such as hardware commoditization, open -source software, virtualization, workforce globalization, and agile IT processes have supported the development of new technology and business models. Cloud computing is a natural evolution of the widespread adoption of virtualization, service-oriented architecture, autonomic and utility computing. In fact, cloud computing is a new term for a long-held dream of computing as a utility, which has recently emerged as a commercial reality. This evolution started in the fifties with mainframe computing allowing that multiple users were able of accessing a central computer through very limited terminals. Later, in the seventies, the concept of virtual machines was created. The development of cloud computing gathered momentum in the nineties, when the Internet started to provide significant bandwidth and telecommunications companies started offering virtualized private network connections.

Some experts attribute the cloud concept to John McCarthy, professor at Stanford University and inventor of Lisp, who proposed in 1961 the idea of computation being delivered as a public utility, similar to the service bureau. In 1966 Douglas F. Parkhill published the book "The Challenge of the Computer Utility" [52] that provides a vision for the future of computing, predicting that the computer industry would come to resemble a public utility "in which many remotely located users are connected via communication links to a central computing facility". Many characteristics of cloud computing (elastic provision, provided as a utility, online, illusion of infinite supply) are included in this book. A. Regalado, in his paper "Who coined "Cloud Computing" [53] states that "many believe the first use of "cloud computing" in its modern context occurred on August 9, 2006, when then Google CEO Eric Schmidt introduced the term to an industry conference. "What's interesting [now] is that there is an emergent new model," Schmidt said, "I don't think people have really understood how big this opportunity really is. It starts with the premise that the data services and architecture should be on servers. We call it cloud computing-they should be in a "cloud" somewhere." One of the first milestones in cloud computing history was the arrival of Salesforce.com in 1999, which pioneered the concept of delivering enterprise applications available from a website. This company paved the way for both specialist and mainstream software firms to deliver applications over the Internet and played a very important role in the introduction of the *Software as a Service (SaaS)*. The SaaS subscription model enables companies to access software online and only pay for the services and applications used. The next development was Amazon Web Services in 2002, which

provided a suite of cloud-based services including storage, computation and even human. Third-party sites could search and display products from the web site of Amazon and add items to Amazon shopping carts. The initial version of AWS in 2002 was focused more on making information available from Amazon to partners through a web services model with programmatic and developer support and was very focused on Amazon as a retailer. In August 2006 Amazon launched as a commercial web service its *Elastic Compute cloud (EC2)*. This solution gives the users a new way to store data offsite, rent compute cycles as a service, and provides online services for other web sites or client-side applications. Probably, EC2 was the first widely accessible cloud computing with infrastructure as a service model.

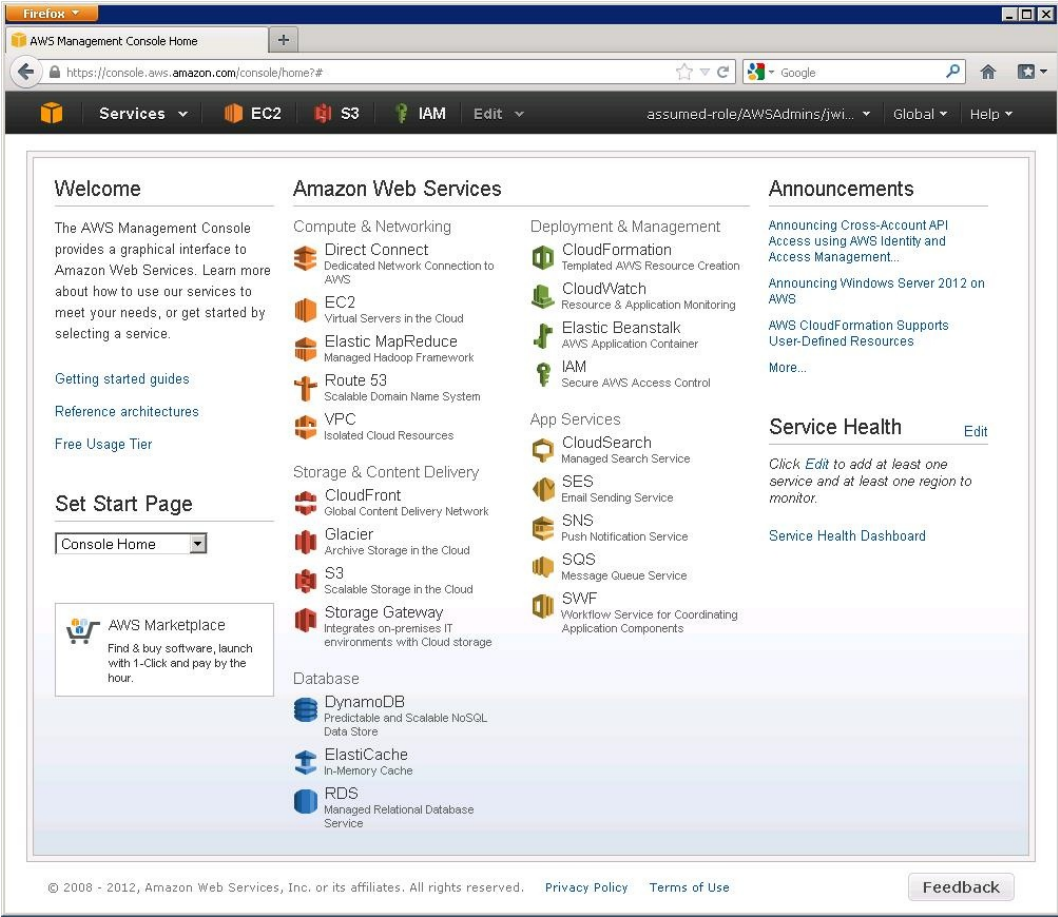


Figure 6.1 - Amazon Web Services Management Console

The launch of Google App Engine in April 2008 was the entry of the first pure play technology company into the Cloud Computing market. Google Apps services allow this company started to offer browser-based enterprise applications. Microsoft for several years did not accept the web as a significant market and continued to focus on the desktop market. Nevertheless, in November 2009, Microsoft changed this criterion and launched its Windows Azure cloud computing platform. This platform provides both PaaS and IaaS service and supports many different programming languages, tools and frameworks. It was renamed to Microsoft Azure in 2014. In December 2013 Google Compute Engine was launched. This infrastructure enables users to create and run virtual machines on demand with a variety of configurations. In 2009-2010, the open source cloud movement gained influence. There are numerous cloud computing services that are either written entirely in open source code, or at least incorporating open source into the final application. The use of open source code in cloud computing allows developers

to build applications on top of an existing application infrastructure, gaining a lower cost, greater flexibility and probably more robust applications (with fewer flaws) than these ones entirely built from scratch.

Across the many cloud computing service models, there are a large and diverse number of applications to choose from and both commercial and free open source offerings. Some examples are Apache CloudStack, Eucalyptus, OpenNebula and OpenStack [54]. The emergence of killer apps from Microsoft, Google, Amazon, Apple, Adobe, Cisco and other big IT companies results in a wider acceptance of online services and constitutes a very relevant contribution for the deployment of cloud computing.

### **6.3 Characteristics of cloud computing**

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. According to NIST, this cloud model is composed of five essential characteristics: On-demand self-service, Broad network access, Resource pooling, Rapid elasticity, Measured Service). Moreover, there are several common characteristics such as scalability, virtualization, service orientation, advanced security, geographic distribution, etc. Next, the five essential characteristics are shortly described:

1. **On-demand self-service.** Cloud computing provides resources, such as server time and network storage, on demand, i.e. when the consumer wants it. Examples of resources include storage, processing, memory, and network bandwidth. The consumer can unilaterally provision computing capabilities automatically. This is possible by self-service and automation. Self-service means that the consumer performs all the actions needed to acquire the service herself. Her request is automatically processed by the cloud infrastructure, without human interaction on the service provider's side. This characteristic implies a high level of planning, since a consumer can request a new resource (i.e. a virtual machine) at any time, and expects to have it working in few minutes. The cloud provider should monitor trends in resource usage and plan for future situations well in advance.
2. **Broad network access.** Capabilities are available over the network and accessed through different client platforms thanks to the use of standard mechanisms. This not only includes the most common devices (laptops, workstations, etc.) but also this includes mobile phones, thin clients and so on. Contrast "broad network access" with access to compute and network resources during the mainframe era. Network, storage and compute resources were scarce and costly several years ago. Over time costs associated with these resources have decreased due to manufacturing scalability and commoditization of associated technologies. As network bandwidth has increased, network access and scalability has increased accordingly. "Broad network access" can be seen both as a trait of cloud computing and as an enabler.
3. **Resource pooling.** The resources of the service provider are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. This concept is a fundamental premise of scalability in the cloud. Multi-tenant environments, where multiple customers share adjacent resources in the cloud with their peers, are the basis of public cloud infrastructures. With multi-tenancy, there is an inherent increase in



operational expenditures, which can be mitigated by certain hardware configurations and software solutions, such as application and server profiles. The resource pooling characteristic provides the feeling of location independence; the customer generally has no control or knowledge over the location of the provided resources. Without resource pooling and multi-tenancy, the economics of cloud computing do not make financial sense.

4. Rapid elasticity. Elasticity is basically a 'rename' of scalability, that is, the ability to add or remove capacity, mostly processing, memory, or both, when this is needed. The rename of the concept is due to the Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time. Most implementations of scalability are based on adding or removing nodes, servers or instances to or from a pool like a cluster or farm. A well-known example is adding a load balancer in front of a farm of web servers that distributes the requests.
5. Measured service indicates that resource usage is monitored, controlled and reported to the consumer, providing visibility and transparency to consumption rates and costs for both the provider and consumer of the utilized service. This is crucial for billing, access control, resource optimization, capacity planning and other tasks.

## 6.4 Cloud computing components and architecture

Many authors emphasize the use and access of multiple server-based computational resources when they refer to cloud computing architecture. Nevertheless, the architecture of a cloud solution is the structure of the system, which typically comprises cloud resources (back end platforms, servers and storage), services, network, middleware, and software components, the externally visible properties of those, and the relationships between them [55].

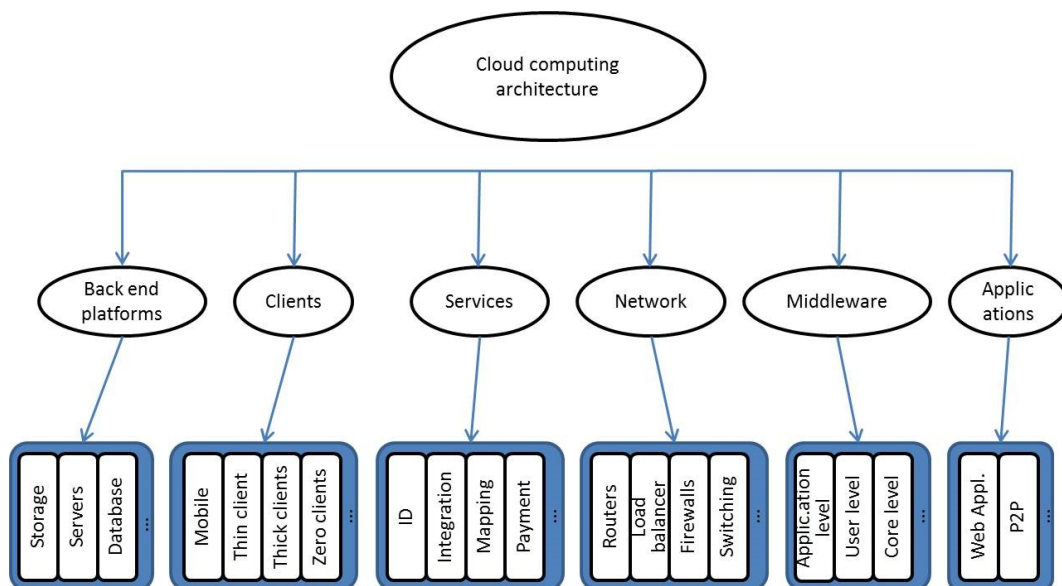


Figure 6.2 - Cloud computing components

Next, the different components are introduced.

1. Back end platforms. These servers are very large, can hold massive amounts of data and can be housed (distributed) anywhere in the world. Often servers are in geographically different places, but they act as if they are working next to each other. Moreover, usually there is a central server that manages the system while at the same time is used for monitoring traffic and client demands to ensure everything properly runs.
2. Cloud users may access the server resources using cloud clients, including fat (or thick) clients, thin clients, zero clients, tablets and mobile devices. These client platforms interact with the cloud data storage via an application (middleware), via a web browser, or through a virtual session.
3. Network: Regarding the client side, the exploitation of cloud computing services by enterprises threatens to take IT server and server management challenges and exchange them for network and network management challenges. Therefore, there are more network requirements. By the other side, regarding the service provider, the network capabilities must ensure that all the communications happen seamlessly, efficiently and in a secure manner. It is critical to have an intelligent, reliable and functional network that provides next generation innovations.
4. Middleware is the software that makes possible the connection between any two clients, servers, databases or even applications. Cloud middleware provides a number of functionalities to the user, helping her in the creation of business applications; facilitating concurrency, transactions, threading and messaging.
5. Services. Cloud services are services that support cloud-based solutions, such as identity management, service-to-service integration, mapping, billing/payment systems, search, ...
6. Applications. A cloud application (or cloud app) is an application program that functions in the cloud, with some characteristics of a pure desktop app and some characteristics of a pure Web app. Usually, these applications are built over a high-level integrated environment; an example is Google's App Engine, which enables users to build Web applications on the same scalable systems that power Google applications.

## 6.5 Service models

Cloud computing offers organizations new choices regarding how to run infrastructures, save costs, and delegate liabilities to third-party providers. It has become an integral part of technology and business models, and has forced businesses to adapt to new technology strategies. Accordingly, the demand for cloud computing has forced the development of new market offerings, representing various cloud service and delivery models. These models significantly expand the range of available options and task organizations with dilemmas over which cloud computing model to employ

Cloud service models describe how cloud services are available to clients. According to NIST, there are three service models: **SaaS** (*Software as a Service*), **PaaS** (*Platform as a Service*) and **IaaS** (*Infrastructure as a Service*) that are described in the following sections. In fact, most fundamental service models include a combination of IaaS, PaaS and SaaS. These service models may have synergies between each other and be interdependent; for example, PaaS is dependent on IaaS because application platforms require physical infrastructure.

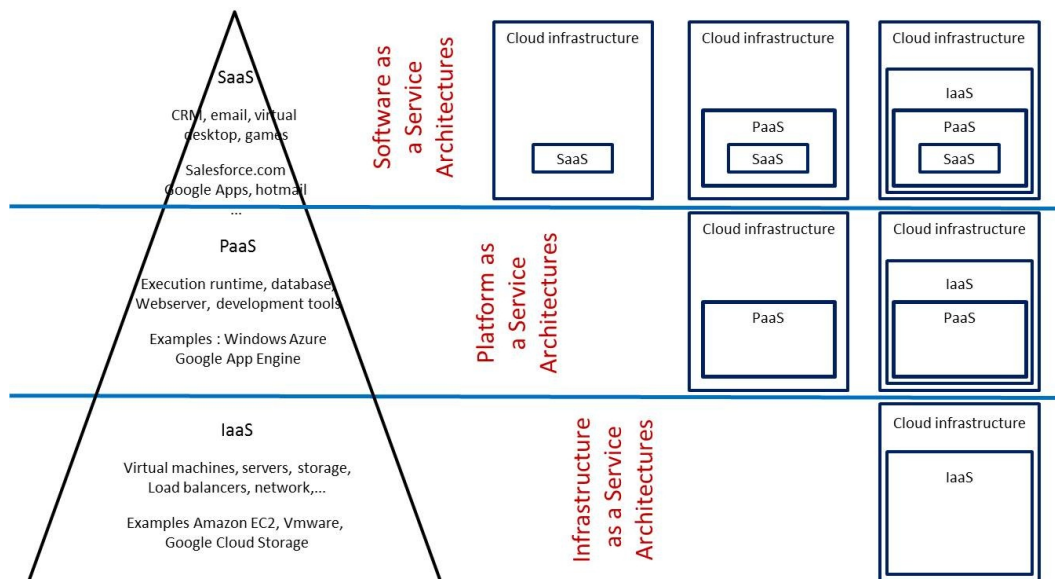


Figure 6.3 - Service models (adapted from [56])

Today, companies realize more value and resource-savings from software and platform services rather than infrastructure. Therefore, IaaS service delivery model is likely to keep losing market share to PaaS and SaaS. It is expected that in the near future significant number of market consolidations with few large players retaining market control at the end [57].

### 6.5.1 Software as a Service

**SaaS** (*Software as a service*) is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet.

This capability eliminates the need to install software on users' computers, and it can be helpful for mobile or transient workforces.

E-mail is a simple example for SaaS.

If a user has a service provider, he requires a desktop or mobile app to access the e-mail; otherwise he can host it on individual servers.

It is important to point out that the user does not manage or control the underlying cloud infrastructure, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. Some benefits of SaaS are:

1. Cost savings: Little or no capital investment
2. Flexible: Offered as an on-demand service
3. Stable: SaaS applications are installed on reputed, protected, and redundant hardware
4. Rapid deployment: Little to no time to provision and deploy
5. Accessible: Only thing needed is an Internet access

6. New releases (upgrade): Service providers upgrade the solution and it becomes available for their customers; the associated costs and efforts are lower than the traditional model

Besides the lack of control, one of the main disadvantages is that SaaS applications may not have the same features as non-SaaS applications. The functionality is often not as refined or as full. Nevertheless, this problem will diminish over time. Development tools for SaaS applications are becoming much more capable. Finally, the speed can be other disadvantage; generally, SaaS applications are slower than the corresponding non SaaS equivalents.

Some examples of SaaS providers are

1. Google Apps: provides web-based office tools such as e-mail, calendar and document management
2. salesforce.com: provides a full *customer relationship management (CRM)* application
3. zoho.com: provides a large suite of web-based applications, mostly for enterprise use

### 6.5.2 Platform as a Service

**PaaS** (*Platform as a service*) is a category of cloud computing that provides a platform and environment to allow developers to build applications and services over the Internet. PaaS services are hosted in the cloud and accessed by users simply via their web browser. Basically, it is a way to rent hardware, operating systems, storage and network capacity over the Internet. The service delivery model allows the customer to rent virtualized servers and associated services for running acquired or consumer-created applications developed using programming languages, libraries, services, and tools supported by the provider. The application development platforms allow users to create and host applications of a larger scale than an individual or small business would be able to handle. PaaS providers can assist developers from the conception of their original ideas to the creation of applications, and through to testing and deployment. This is all achieved in a managed mechanism.

The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment. They benefit from the economies of scale that arise from the sharing of the underlying physical infrastructure between users, and that results in lower costs; PaaS services are generally paid for on a subscription basis with clients ultimately paying just for what they use.

Some examples of the features that can be included with a PaaS are:

1. Operating system
2. Server-side scripting environment
3. Database management system
4. Server Software
5. Support
6. Storage
7. Network access

8. Tools for design and development
9. Hosting

Some benefits of PaaS are:

1. Regarding software developers, they can use individual PaaS environments at every stage of the process to develop, test and ultimately host their applications.
2. Teams can collaborate. Anyone in any location has the ability to work on software projects.
3. Flexibility; customers can have control over the tools that are installed within their platforms and can create a platform that suits their specific requirements.
4. Cost savings: there is no need to invest in physical infrastructure
5. Maximize uptime: PaaS vendors should have the tools, technologies, and experience to help the user to avoid the unplanned outages that cause downtime
6. Scale easily. Features can be changed if circumstances dictate that they should.

One of the cons of PaaS is that depending on the offerings of the company providing the PaaS, the user could be locked into a specific software environment, language or interface. This can affect some, not all, providers.

Some PaaS examples are:

1. Google App Engine: provides users with a complete development stack and allows them to run their applications on Google's infrastructure
2. Akamai EdgePlatform: provides a large distributed computing platform on which organizations can deploy their web applications; has a large focus on analysis and monitoring
3. Microsoft Azure Services Platform: provides users with on-demand compute and storage services as well as a development platform based on Windows Azure
4. *Yahoo! Open Strategy (Y!OS)*: provides users with a means of developing web applications on top of the existing Yahoo! Platform and in doing so leveraging a significant portion of the Yahoo! Resources

### 6.5.3 Infrastructure as a Service

**IaaS** (*Infrastructure as a service*) is a provision model in which an organization outsources the underlying operating systems, security, networking, storage and servers for developing such applications, services, and for deploying development tools, databases, etc. The service provider owns the equipment and is responsible for housing, running and maintaining it. The client typically pays on a per-use basis. Instead of ready-made applications or services, just network is provided; some of the most common uses of IaaS include virtual servers, load balancers, and network connections. IaaS allows organizations and developers to extend their IT infrastructure on an on-demand basis. The cloud provider has pool of virtualized computing resources and storage which the customer organization can take advantage of. This is on demand computing and takes care of the variation in computing peaks. Physically, the pool of

hardware resource is pulled from a multitude of servers and networks usually distributed across numerous data centres, all of which the cloud provider is responsible for maintaining.

The client, on the other hand, is given access to the virtualised components in order to build their own IT platforms. She does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g. host firewalls). Some benefits of IaaS are:

1. Quick and easy access to enterprise class capabilities,
2. Scalability; resource is available as and when the client needs it and, therefore, there are no delays in expanding capacity or the wastage of unused capacity
3. Simplicity: the provider assumes the facilities management, hardware/software procurement, provisioning, patching, and all the other complex details involved with infrastructure.
4. No investment in hardware; the underlying physical hardware that supports an IaaS service is set up and maintained by the cloud provider, saving the time and cost of doing so on the client side
5. Location independence; the service can usually be accessed from any location as long as there is an internet connection and the security protocol of the cloud allows it
6. Physical security of data centre locations; services available through a public cloud, or private clouds hosted externally with the cloud provider, benefit from the physical security afforded to the servers which are hosted within a data centre
7. Rapid deployment: Little to no time to provision and deploy.

In bigger businesses, the main advantage is the last one; it is more related with timely deployment in supporting short term and unforeseen needs.

The main disadvantage of the IaaS is the business risk. Even with extensive diligence, ongoing audits and proactive management, IaaS still requires trust in the vendor infrastructure/operations for availability, data security etc.

Some examples of IaaS providers are

1. Amazon *Elastic Compute Cloud (EC2)*: provides users with a special virtual machine (AMI) that can be deployed and run on the EC2 infrastructure
2. Amazon *Simple Storage Solution (S3)*: provides users with access to dynamically scalable storage resources
3. Microsoft Live Mesh: provides users with access to a distributed file system; targeted at individual use
4. IBM *Computing on Demand (CoD)*: provides users with access to highly configurable servers plus value-added services such as data storage

In common with the other two forms of cloud hosting, IaaS can be utilised by enterprise customers to create cost effective and easily scalable IT solutions where the complexities and expenses of managing the underlying hardware are outsourced to the cloud provider. If the scale of a business customer's operations fluctuates, or they are looking to expand, they can tap into

the cloud resource as and when they need it rather than purchase, install and integrate hardware themselves.

## 6.6 Deployment models

There are four commonly-used cloud deployment models: private, public, hybrid and community. The last one is less-commonly used.

1. A private cloud is built and managed within a single organization. Organizations use software that enables cloud functionality, such as VMWare.
2. A public cloud is a set of computing resources provided by third-party organizations. The most popular public clouds include Amazon Web Services, Google AppEngine, and Microsoft Azure.
3. A hybrid cloud is a mix of computing resources provided by both private and public clouds.
4. A community cloud shares computing resources across several organizations, and can be managed by either organizational IT resources or third-party providers

Public cloud deployment model is likely to stay dominant and keep expanding further. Private and Hybrid deployment models are going to stay for years ahead but their market share is going to continuously drop. In the long-term private and hybrid cloud models most probably will be used only for specific business cases.

### 6.6.1 Public cloud

Public clouds are made available to the general public by a service provider who hosts the cloud infrastructure. Generally, public cloud providers like Amazon AWS, Microsoft and Google own and operate the infrastructure and offer access over the Internet for open use by the general public. With this model, customers have no visibility or control over where the infrastructure is located. It is important to note that all customers on public clouds share the same infrastructure pool with limited configuration, security protections and availability variances. Some examples include services aimed at the general public, such as online photo storage services, e-mail services, or social networking sites. However, services for enterprises can also be offered in a public cloud.

In public clouds, resources are offered as a service. Users can scale their use on demand and do not need to purchase hardware to use the service. Public cloud providers manage the infrastructure and pool resources into the capacity required by its users. Public Cloud customers benefit from economies of scale, because infrastructure costs are spread across all users, allowing each individual client to operate on a low-cost, "pay-as-you-go" model. Another advantage of public cloud infrastructures is that they are typically larger in scale than an in-house enterprise cloud, which provides clients with seamless, on-demand scalability. These clouds offer the greatest level of efficiency in shared resources.

A public cloud is the obvious choice when:

1. The standardized workload for applications is used by lots of people, such as e-mail.
2. It is needed to test and develop application code.

3. An incremental capacity is required (the ability to add compute resources for peak times).
4. Collaboration projects are being done.

### 6.6.2 Private cloud

In a private cloud, the cloud infrastructure is a proprietary operated solely for a single client or organization. It is not shared with other organizations, but it can comprise multiple consumers (e.g. business units). The infrastructure may be hosted internally or externally, and it can be managed by the organization, by a third party or by some combination of them. It allows organizations to host data and applications in the cloud, in a more secure and controlled environment regarding public clouds. The resources are deployed behind a firewall and may be accessed across private leased lines or secure encrypted connections via public networks to get the specified clients (and only they) can operate. The objective of these mechanisms is to minimize security concerns and limit access to specific clients. There are two variations of private clouds according to the place where it is hosted [58]:

1. **On-Premise Private Cloud:** This type of cloud is hosted within an organization's own facility. This option is suitable for organizations that have invested in substantial server and storage hardware and want to leverage that investment and repurpose some of such equipment in a private cloud, either for applications that require complete control and configurability of the infrastructure and security.
2. **Externally Hosted Private Cloud:** Externally hosted private clouds are hosted by a third party specializing in cloud infrastructure. The service provider facilitates an exclusive cloud environment with full guarantee of privacy. This format is recommended for organizations that prefer not to use a public cloud infrastructure due to the risks associated with the sharing of physical resources.

Undertaking a private cloud project requires a significant level and degree of engagement to virtualize the business environment. Private clouds are more expensive but also more secure when compared to public clouds. An important part of IT decision-makers will focus exclusively on the private cloud, as these clouds offer the greatest level of security and control.

As a summary, a private cloud is the best option when

1. A high control level is needed.
2. Data security and privacy are critical.
3. Data sovereignty is required but cloud efficiencies are desired

### 6.6.3 Community cloud

A community cloud is a multi-tenant cloud service model that is shared among several organizations from a specific group with common computing concerns (e.g., mission, security requirements and compliance considerations). These organizations or communities have similar cloud requirements and their ultimate goal is to work together to achieve their business objectives. Community Cloud has its own challenges like allocation of costs, responsibilities, governance and security. The costs are spread over fewer users than a public cloud (but more than a private cloud), so only some of the cost savings potential of cloud computing are realized.



The cloud may be managed by the organisations or by a third party and can be either on premise or off-premise.

In general, public cloud services are likely to be more cost efficient and scalable than private clouds, but less secure. The goal of community clouds is to have participating organizations realize the benefits of a public cloud with the added level of privacy, security, and policy compliance usually associated with a private cloud.

Government, healthcare, telco community, and some regulated private industries are leveraging the added security features within a community cloud environment. Instead of just provisioning space in a public cloud, organizations can test and work on a cloud platform which is secure, "dedicated," and even compliant with certain regulations.

#### 6.6.4 Hybrid cloud

Hybrid Cloud or combined Cloud is a combination of different methods of resource pooling (for example, combining public and community clouds). Also, hybrid clouds can be seen as a composition of two or more clouds (private, community or public) that are bound together, offering the advantages of multiple deployment models, by standardized or proprietary technology that enables data and application portability. These entities can leverage third party cloud providers in either a full or partial manner; increasing the flexibility of computing. The idea of hybrid cloud is to combine several cloud models to create a customized solution based on the organization requirements. Hybrid cloud architecture requires both on premise resources and off-site server based cloud infrastructure and can be implemented in a number of ways. For instance, it is possible for an entity to have its data and applications in cloud maintaining control over organizational network topology and policies, and at the same time keeping its existing physical infrastructure (although this infrastructure does not scale) and borrow additional resources when required.

Hybrid clouds allow a company to keep each aspect of its business in the most efficient environment possible. Moreover, these models are easily scalable, cost efficient (it is possible to reduce the demand on a private cloud by moving non-sensitive data or applications to the public cloud), secure and flexible.

The downside is that they have to keep track of multiple cloud security platforms and ensure that all aspects of the business can communicate with each other.

Hybrid cloud is a good option when:

1. A company might use a public cloud for test and development while using a hosted private cloud inside the organization for production deployment.
2. A company that uses public clouds for external facing applications while using a hosted private cloud for internal applications.
3. A company that wants to use a public cloud to interact with the customers but keep their sensitive data secured within a private cloud.

### 6.7 Uses and applications

Cloud computing can support nearly any application. But some workloads are a better fit for cloud from an organization or technical perspective. The two killer use cases for cloud computing are:

1. Managing big data. Cloud computing is a natural fit for big data analytics, since elastic computing capacity and on-demand provisioning make analytics accessible to more teams within an organization. Moreover, cloud is also a useful solution when lots of computations are required for solving complex problems, or when collaboration among developers is needed. Cloud computing allows much more efficient computation by centralizing storage memory processing and bandwidth
2. Test and developments in the cloud. Development teams will benefit from the agility of creating virtual machines in minutes. Cloud computing avoids that organizations need to set up an environment through physical assets, significant manpower and time. Moreover, an organization can quickly get applications into production and to scale them as required

Other relevant uses are:

1. File storage and sharing
2. Backup and disaster recovery
3. CRM
4. Web site hosting

Typical applications are:

1. Social networks sites
2. E-mail sites
3. Search engines
4. Communications facilities (i.e. Skype)
5. Time-tracking applications
6. Notes organizer (like Evernote)
7. Creating and sharing office documents (google apps)

## **6.8 Benefits and disadvantages of cloud computing**

Many users, businesses large and small use cloud computing today either directly (e.g. Google or Amazon) or indirectly (e.g. Twitter) instead of traditional alternatives. Clouds can provide users with a number of different benefits. One of the most relevant is the reduction of cost and complexity of owning and operating computers and networks. Cloud users do not need to invest in information technology infrastructure, purchase hardware, or buy software licences. Moreover, there are clouds providers specialized in particular areas (such as e-mail) that can bring advanced services very useful for a single company. Some other benefits to clients include scalability, reliability, and efficiency. Scalability means that cloud computing offers unlimited processing and storage capacity. The cloud is reliable in that it enables access to applications

and documents anywhere in the world via the Internet. Cloud computing is often considered efficient because it allows organizations to free up resources to focus on innovation and product development. In addition, information in the cloud is not as easily lost.

Next, there is a list of some of the most important benefits derived of the use of cloud computing:

1. Availability and universal access. Cloud computing can allow remotely located employees to access resources and applications resources at any time through a standard internet connection.
2. Choice of applications. This allows flexibility for cloud users to experiment and choose the best option for their needs. Cloud computing also allows a business to use, access and pay only for what they use, with a fast implementation time
3. Collaboration. Users begin to see the cloud as a way to work simultaneously on common data and information.
4. Cost reduction. The pay-per-usage model, unlike on-site hosting, allows an organization to only pay for the resources they need with basically no investment in the physical resources available in the cloud.
5. Elasticity. The provider transparently manages the clients resource utilization based on dynamically changing needs.
6. Flexibility. Cloud computing allows clients to switch applications easily and rapidly, using the one that suits their needs best.
7. Potential to be greener and more economical. The average amount of energy needed for a computational action carried out in the cloud is far less than the average amount for an on-site deployment. This is because different organisations can share the same physical resources.
8. Risk reduction. Organizations can use the cloud to test ideas and concepts before making major investments in technology.
9. Scalability. Users have access to a large amount of resources that scale based on their demand.
10. Up to date software. A cloud provider will also be able to upgrade software keeping in mind feedback from previous software releases.
11. Virtualization. Each user has a single view of the available resources, independently of how they are arranged in terms of physical devices. Therefore, there is potential from a provider perspective to serve a greater number of users with fewer physical resources.

Nevertheless, there are some problems that can act as a barrier when an organization wants to adopt cloud computing. Following there is a list of these concerns:

1. Interoperability. A universal set of standards and/or interfaces have not yet been defined, resulting in a significant risk of vendor lock-in.
2. Latency. All access to the cloud is done via the internet, introducing latency into every communication between the user and the provider.

3. Platform or Language Constraints. Some cloud providers support specific platforms and languages only.
4. Regulations. There are concerns in the cloud computing community over jurisdiction, data protection, fair information practices, and international data transfer, mainly for organizations that manage sensitive data.
5. Reliability. Many existing cloud infrastructures leverage commodity hardware that is known to fail unexpectedly.
6. Resource Control. The amount of control that the user has over the cloud provider and its resources vary greatly between providers.
7. Security. The main concern is data privacy: users do not have control or knowledge of where their data is being stored. Nevertheless, regarding forensic security, the use of cloud computing (when using virtualisation) can provide dedicated, pay-per-use forensic images of virtual machines which are accessible without taking infrastructure off-line, leading to less down-time for forensic analysis. It can also provide more cost-effective storage for logs allowing more comprehensive logging without compromising performance [59].

## **6.9 Cloud security (potential privacy risks)**

While there are benefits, there are privacy and security concerns too. Data is travelling over the Internet and is stored in remote locations. In addition, cloud providers often serve multiple customers simultaneously. All of this may raise the scale of exposure to possible breaches, both accidental and deliberate. It is needed to ensure that the personal information is appropriately handled [60]. Security concerns may be magnified by the dynamic nature of the cloud environment. One of the cloud key advantages is the speed with which the cloud vendors can adjust, develop and change their offerings. There is a trade-off between this speed and flexibility requirements and the security level. Cloud computing poses several data protection risks for cloud customers and providers. In some cases, it may be difficult for the cloud customer (in its role as data controller) to effectively check the data handling practices of the cloud provider and thus to be sure that the data is handled in a lawful way. This problem increases in cases of multiple transfers of data, e.g. between federated clouds. Privacy, including the need to protect identity information, is a core issue in the success of cloud computing deployment. Many organizations do not feel comfortable storing their data and applications on systems that reside outside of their on-premise data centres. The risk of exposure or unauthorized access of sensitive data increases when workloads migrate to a shared infrastructure. Concerns have been raised by many that cloud computing may lead to "function creep" - uses of data by cloud providers that were not anticipated when the information was originally collected and for which consent has typically not been obtained. When a request to delete a cloud resource is made, as with most operating systems, this may not result in true wiping of the data. Adequate or timely data deletion may also be impossible (or undesirable from a customer perspective), either because extra copies of data are stored but are not available, or because the disk to be destroyed also stores data from other clients. In the case of multiple tenancies and the reuse of hardware resources, this represents a higher risk to the customer than with dedicated hardware. Given how inexpensive it is to keep data, there is little incentive to remove the information from the cloud and more reasons to find other things to do with it

Cloud service providers must assure their customers and provide a high degree of transparency into their operations and privacy assurance. Privacy-protection mechanisms must be embedded in all security solutions.

## 6.10 Conclusions

Cloud computing is a nascent and rapidly evolving model, with new aspects and capabilities being announced regularly. Specifically, cloud computing usually refers to a cloud alternative to something that organizations would traditionally manage in-house, using dynamically scalable and often virtualized resources provided as a service over the Internet/Intranet. For example, a webmail service is a cloud-based alternative to hosting your own email server. Most cloud computing services are accessed through a web browser using any connected device (mobile, tablet, personal computer, ...). Therefore, cloud services do not require users to have sophisticated computers that can run specialized software. A user-centric interface makes the cloud infrastructure supporting the applications transparent to users. Cloud computing has the potential to be a disruptive force by affecting the deployment and use of technology; in fact, cloud is changing the way that many organizations manage information technology. Depending on the perspective and situation of the organization or the individual, this represents both opportunity and crisis. Such change may be resisted, even if it is a good idea and it works. Companies have a range of paths to the cloud, including infrastructure, platforms and applications that are available from cloud providers as online services. One of the main concerns is security and privacy. This concern depends on the kind of company. In the case of big organizations with significant resources to devote to a sophisticated information security program, it is needed to overcome a number of security, privacy, and compliance challenges. Nevertheless, in the case of a *small to medium-size business (SMB)*, the security of cloud computing might look attractive, compared to the resources the company can afford to spend on information security today.

## 7. Network virtualisation - SDN&NFV

### 7.1 SDN

The traditional packet network architecture consists of nodes integrating two essential components - packet control logic and packet forwarding hardware (*forwarding* is a term used for deciding the output port or set of output ports for a packet, and transferring that packet to those output ports). In most of routers (or any networking equipment for that matter) there is specialized hardware for fast forwarding of data between interfaces - it creates so-called Data Forwarding plane. The forwarding is managed by rules created by processor running operating system, routing algorithms, address translation, and other higher functions - it creates so-called Control plane.

Figure 7.1 shows the architecture of a traditional packet network in which each network node contains a data and control plane. This model limits the introduction of new features and protocols (each new feature or protocol must be implemented in each network node) and brings the complexity of control mechanisms and limits innovation in IP networks. Extreme growing number of connected devices also can limit future addressing space and usability/scalability of existing routing protocols for large complex networks, as well as further innovation.

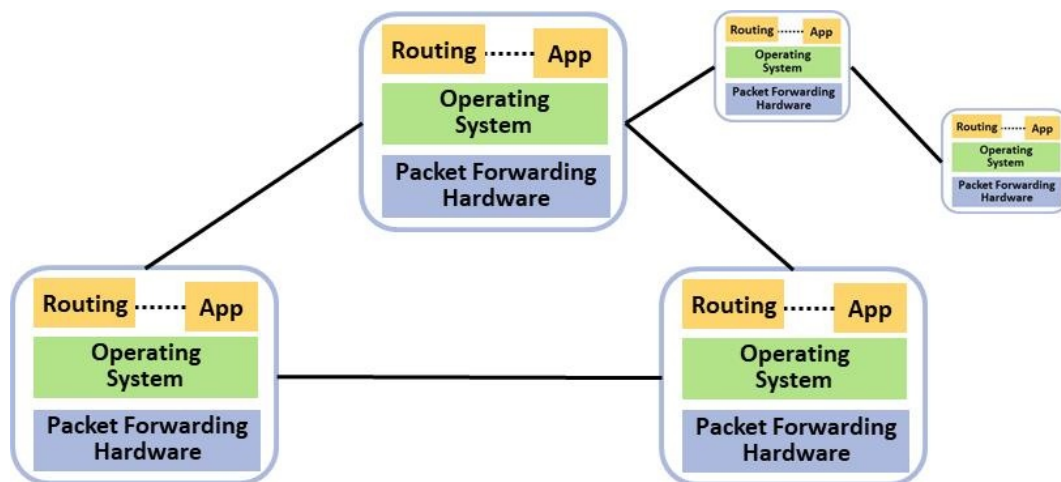


Figure 7.1 - Traditional packet network architecture

*Software-Defined Networking (SDN)* is an emerging ICT architecture that decouples the network control and forwarding functions enabling the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services.

The SDN defines separation of the control and data forwarding plane in network as depicted in Figure 7.2. By the implementing of the separated control plane by the software for general purpose computer (called Network OS) from forwarding plane on network equipment, it is possible to centralize routing and switching decisions, as well as configuration of all network devices. SDN allows network administrators to automate and dynamically manage and manage a large number of network devices, services, topology, data paths, and packet processing (QoS) policies through higher-level languages and the *Application Programming Interface (API)*.

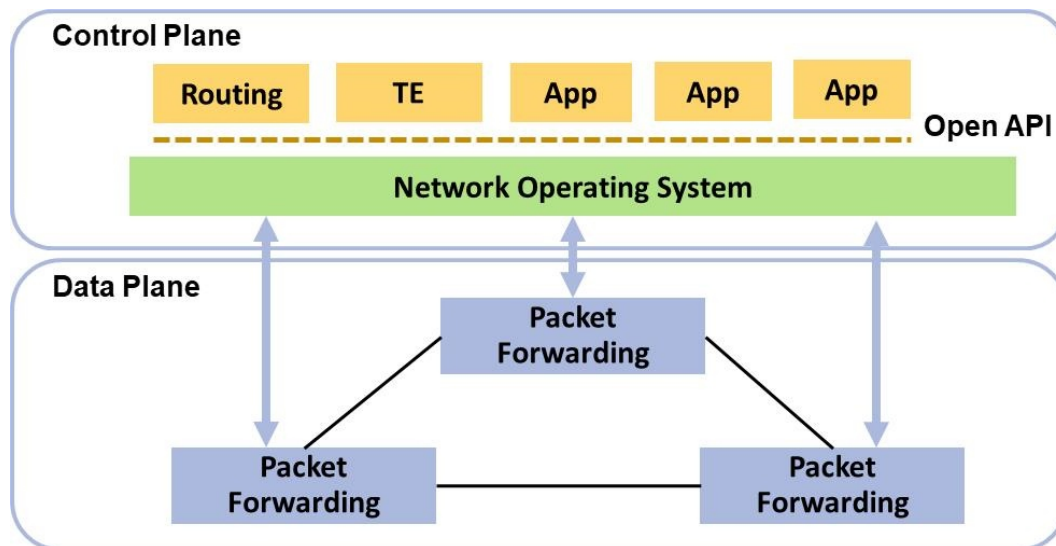


Figure 7.2 - Basic SDN architecture

The basic idea of SDN is the separation of the control plane from the data forwarding plane in the network. Removing the management plane from each network node and using a centralized management plane will not only allow a complete view of the network without convergence problems that are natural to the distributed algorithms, but also reduce costs. Additionally, without the need to distribute routing functions to individual network nodes, it is much easier to introduce new ones, or to modify and improve existing algorithms. In addition, configuring a central management plane makes network management easier, reduces the possibility of misconfiguration, and speeds up troubleshooting. An additional advantage of implementing software in a centralized management plane is the easy modification and development of new features.

SDN based network can enable:

- Simpler architecture - separation of control from forwarding data plane, potentially integration of control plane of IP & transport networks
- Resilience and Automation - self-healing mechanisms, better scale in/out, higher level of automation, end-to-end network visibility
- Time to market - faster delivery of new service, hosting network functions in infrastructure cloud

## Open Flow

**OpenFlow** is an open standard originally developed at universities and currently maintained by *Open Network Foundation (ONF)* - a non-profit consortium with mission to commercialize and promote OpenFlow based SDN. OpenFlow is the most popular protocol used for communication between control plane and data forwarding plane - becoming the de facto standard.

The core of the OpenFlow specification is a switch (i.e. packet processing equipment). The switch processes packets according the packet content and the configured status of the switch. The switch communicates with the controller and the controller manages the switch via the

OpenFlow switch protocol. A controller as the main part of the control plane uses the OpenFlow protocol to communicate with multiple switches and to control configuration of their statuses.

In general, the OpenFlow specifications describe:

- the southbound protocol between OpenFlow controller and OpenFlow switches
- the operation of the OpenFlow switch

### OpenFlow Switch

The OpenFlow switch is an integral part of the OpenFlow interface. There are two types of OpenFlow-compliant switches:

- ***OpenFlow only switches*** - support only OpenFlow operation, in those switches all packets are processed by the OpenFlow pipeline, and cannot be processed otherwise.
- ***OpenFlow-hybrid switches*** - support both OpenFlow operation and normal Ethernet switching operation. An OpenFlow-hybrid switch may also allow a packet to go from the OpenFlow pipeline to the normal pipeline through the NORMAL and FLOOD reserved ports.

The OpenFlow switch uses its output port to forward packets to the input port of another switch or device. Only standard ports, such as physical ports, logical ports, and reserved ports (if supported) can be used as input and output ports. The reserved ports are both mandatory (required) and non-mandatory (optional) and are defined in the OpenFlow specification [61].

The OpenFlow switch consists of several parts (see Figure 7.3) needed to process the packet. It contains one or more *flow tables* with packet processing rules and *group table* that together serve for packet search and processing. Each rule belongs to a certain subset of data flows and performs operations for that flows. The operations performed include packet dropping (discarding), packet forwarding, or packet forwarding to all available ports (so-called flooding).

According the rules introduced by the controller, the OpenFlow switch can act as a router, switch, firewall, network address translator, or something in between.



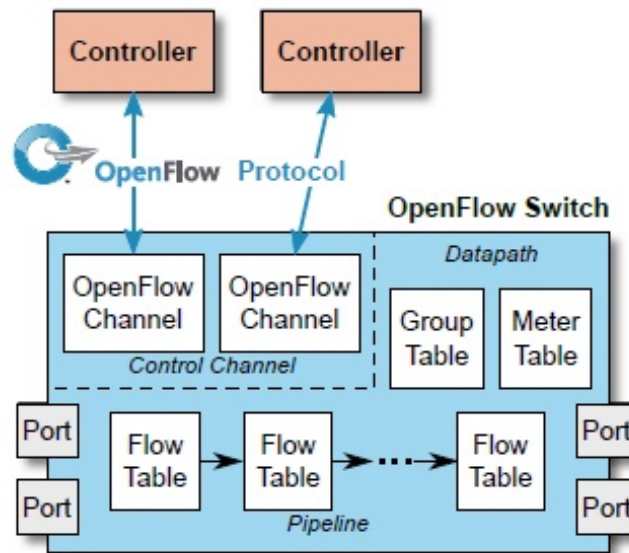


Figure 7.3 - Main components of an OpenFlow switch [61]

The OpenFlow Switch consists of one or more *flow tables* and a *group table*, which perform packet lookups and forwarding, and one or more *OpenFlow channels* to an external controller (Figure 7.3). Using the OpenFlow switch protocol, the controller can add, update, and delete *flow entries* in flow tables, both reactively (in response to packets) and proactively. Each flow table in the switch contains a set of flow entries; each flow entry consists of *match fields*, *counters*, and a set of *instructions* to apply to matching packets.

Matching starts at the first *flow table* and may continue to additional flow tables of the pipeline. Flow entries match packets in priority order, with the first matching entry in each table being used. If a matching entry is found, the instructions associated with the specific flow entry are executed (see Figure 7.4). If no match is found in a flow table, the outcome depends on configuration of the table-miss flow entry: for example, the packet may be forwarded to the controllers over the OpenFlow channel, dropped, or may continue to the next flow table. (Every flow table must support a table-miss flow entry to process table misses).

Instructions associated with each flow entry either contain actions or modify pipeline processing. Actions included in instructions describe packet forwarding, packet modification and group table processing. Pipeline processing instructions allow packets to be sent to subsequent tables for further processing and allow information, in the form of metadata, to be communicated between tables. Table pipeline processing stops when the instruction set associated with a matching flow entry does not specify a next table; at this point the packet is usually modified and forwarded.

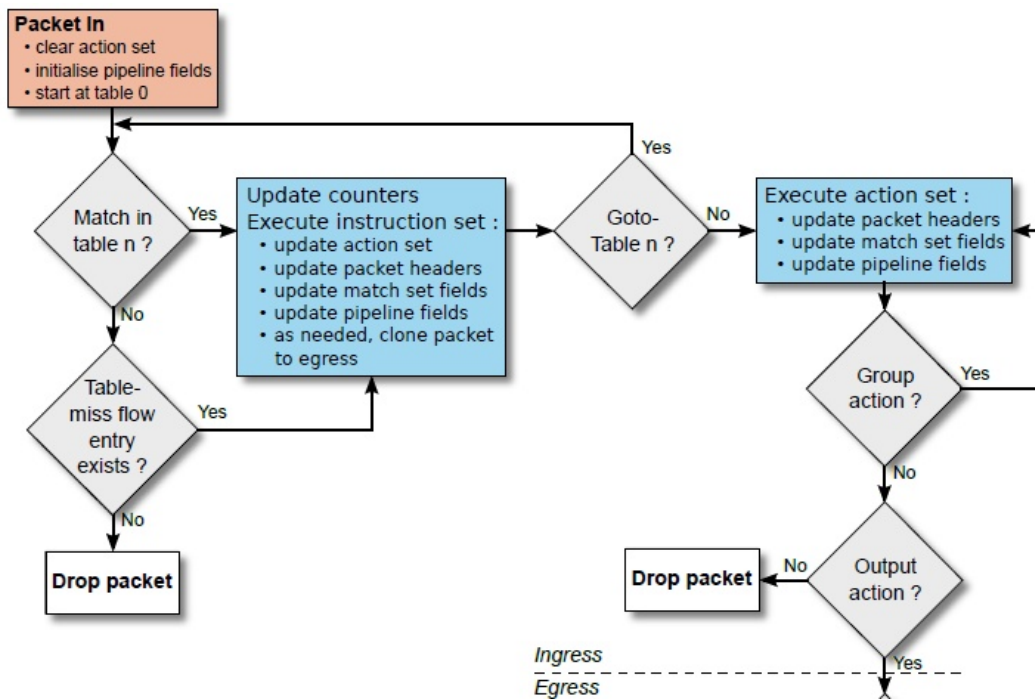


Figure 7.4 - Simplified flowchart detailing packet flow through an OpenFlow switch [61]

Flow entries may forward to a port. This is usually a physical port, but it may also be a logical port defined by the switch or a reserved port defined by this specification. Reserved ports may specify generic forwarding actions such as sending to the controller, flooding, or forwarding using non-OpenFlow methods, such as “normal” switch processing, while switch-defined logical ports may specify link aggregation groups, tunnels or loopback interfaces.

Actions associated with flow entries may also direct packets to a group, which specifies additional processing. Groups represent sets of actions for flooding, as well as more complex forwarding semantics (e.g. multipath, fast reroute, and link aggregation).

The **group table** contains group entries; each group entry contains a list of action buckets with specific semantics dependent on group type. The actions in one or more action buckets are applied to packets sent to the group.

A **meter table** consists of meter entries, defining per-flow meters. Per-flow meters enable OpenFlow to implement rate-limiting, a simple QoS operation constraining a set of flows to a chosen bandwidth. Per-flow meters can also enable OpenFlow to implement more complex QoS policing operations. A meter measures the rate of packets assigned to it and enables controlling the rate of those packets. Meters are attached directly to flow entries (as opposed to queues which are attached to ports).

If the switch does not find a suitable record for the packet in its tables, it can send the packet by the PACKET-IN message to the control unit. The controller can send instructions by the PACKET-OUT message to complete packet processing in the switch and by the FLOW-MOD message send a command to create a new flow record in some table on the switch.

The OpenFlow switch protocol supports three message types:

- **Controller-to-switch messages** - are initiated by the controller and used to directly manage or inspect the state of the switch.

- **Asynchronous messages** - are initiated by the switch and used to update the controller about network events and changes to the switch state.
- **Symmetric messages** - are initiated by either the switch or the controller and sent without solicitation

Controller-to-switch messages:

- **FEATURES** - The feature request is sent to the switch to request identity and the basic capabilities of the switch. The switch must respond with a features reply that specifies the identity and basic capabilities of the switch.
- **CONFIGURATION** - The controller is able to set and query configuration parameters in the switch. (The switch only responds to a query from the controller.)
- **MODIFY-STATE** - Used to manage state on the switches. (Its primary purpose is to add, delete and modify flow/group entries and insert/remove action buckets of group in the OpenFlow tables and to set switch port properties.
- **READ-STATE** - Used to collect information from the switch (such as current configuration, statistics and capabilities.)
- **PACKET-OUT** - Used to send packets out of a specified port on the switch, and to forward packets received via Packet-in messages. Packet-out messages must contain:
  - A full packet or a buffer ID referencing a packet stored in the switch.
  - A list of actions to be applied in the order they are specified (an empty list of actions drops the packet).
- **BARRIER** - Barrier request/reply messages are used by the controller to ensure message dependencies have been met or to receive notifications for completed operations.
- **ROLE-REQUEST** - Used to set the role of its OpenFlow channel, set its Controller ID, or query these. (This is mostly useful when the switch connects to multiple controllers)
- **ASYNCHRONOUS-CONFIGURATION** - Used to set an additional filter on the asynchronous messages that it wants to receive on its OpenFlow channel.

Asynchronous messages:

- **PACKET-IN** - Transfer the control of a packet to the controller. If the switch is configured to buffer packets and the switch has sufficient memory to buffer them, the PACKET-IN message contains only some fraction of the packet header and a buffer ID. Switches that do not support internal buffering or have run out of internal buffering, must send the full packet to controllers.
- **FLOW-REMOVED** - Inform the controller about the removal of a flow entry from a flow table.
- **PORT-STATUS** - Inform the controller of a change on a port (e.g. if the link went down).

- **ROLE-STATUS** - Inform the controller of a change of its role (e.g. when a new controller elects itself master, the switch send role-status message to the former master controller)
- **CONTROLLER-STATUS** - Inform the controller when the status of an OpenFlow channel changes. The switch sends these messages to all controllers when the status of the OpenFlow channel to any switch changes. (This can assist failover processing if controllers lose the ability to communicate among themselves.)
- **FLOW-MONITOR** - Inform the controller of a change in a flow table. (A controller may define a set of monitors to track changes in flow tables)

Symmetric messages:

- **HELLO** - Initialization message exchanged between the switch and controller upon connection startup.
- **ECHO** - Echo request/reply messages are mainly used to verify the liveness of a controller-switch connection. They may be used to measure latency or bandwidth. (*The receiving site must always send a reply*).
- **ERROR** - Used by the switch or the controller to notify problems to the other side of the connection. They are mostly used by the switch to indicate a failure of a request initiated by the controller.
- **EXPERIMENTER** - Intended for experimental purposes, respectively implementing new features.

The switch may establish communication with a single controller, or with multiple controllers. Having multiple controllers improves reliability. The hand-over between controllers is initiated by the controllers themselves, which enables fast recovery from failure and also controller load balancing. When OpenFlow operation is initiated, the switch must connect to all controllers it is configured with, and try to maintain connectivity with all of them concurrently. If any controller sends controller-to-switch command to the switch, the reply or error message related to those command must only be sent on the controller connection associated with that command. Asynchronous messages may need to be sent to multiple controllers.

## 7.2 NFV

*Network functions virtualisation* (NFV) is a network architecture concept that uses the technologies of IT virtualisation to virtualise entire classes of network node functions into building blocks that may connect, or chain together, to create communication services.

Virtualisation means that a network function and part of the infrastructure are implemented in software and therefore the NFV software architecture is an important aspect of the NFV architectural framework.

The first draft of NFV was published in white paper [62] driven mainly by service providers expectations, later the standardization activities went to *European Telecommunications Standards Institute* (ETSI) that created a *Network Functions Virtualisation Industry Specification Group* (ETSI NFV ISG) covering all tasks related to NFV technology.

### 7.2.1 NFV architecture

The architecture of NFV technology is based on the following components (Figure 7.5):

- **NFVI (NFV Infrastructure)** - provides virtual resources needed to support the implementation of virtualised network functions - commercial off-the-shelf hardware components for acceleration, layer of software that virtualises and abstracts the underlying hardware.
- **VNF (Virtualised Network Functions)** - software implementation of network functions that is able to run by NFVI and may be accompanied by **EMS (Element Management System)**, which manages the VNF. VNF is an entity corresponding to today's network node, which is expected to be delivered as pure software independent of the hardware.
- **NFV MANO (Management and Orchestration)** - covers orchestration and lifecycle management of physical and/or software tools that support the virtualisation and infrastructure lifecycle management VNFs. NFV MANO focuses on virtualisation management tasks, which is necessary for NFV framework. It also collaborates with external NFV **OSS / BSS (Operations Support System / Business Support System)** and enables integration NFV to existing networks.

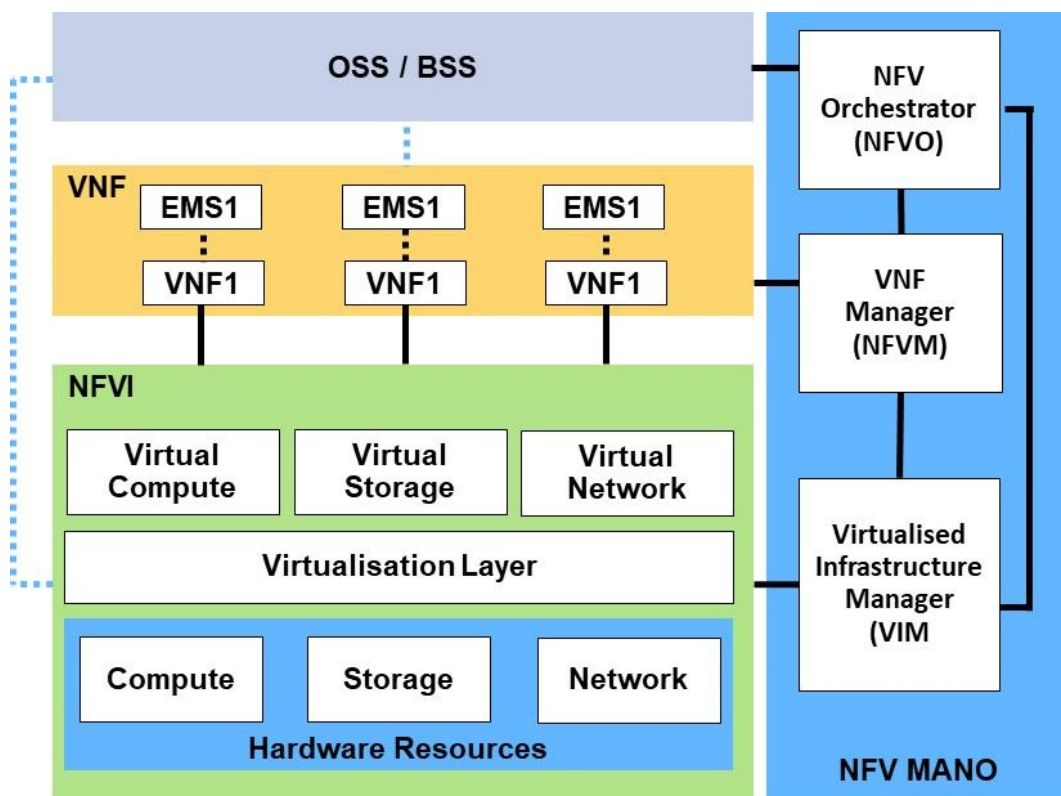


Figure 7.5 - NFV architecture

### 7.2.2 NFV infrastructure

The NFV infrastructure is a sum of all hardware and software components which build up the environment in which NFVs are deployed, managed and executed. The NFV infrastructure can

span across several locations. The network providing connectivity between these locations is regarded to be part of the NFV infrastructure.

The hardware resources include computing, storage and network that provide processing, storage and connectivity to VNFs through the virtualisation layer (e.g. hypervisor) [63], [64]. Computing hardware is assumed to be commercial of-the-shelf hardware, mostly servers in service provider datacentres. Storage resources can be differentiated between shared **NAS** (*Network Attached Storage*) and storage that reside on the server itself.

The virtualisation layer abstracts the hardware resources and decouples the VNF software from the underlying hardware and therefore, the software can be deployed on different physical hardware resources. The primary tools to realize the virtualisation layer would be the hypervisors.

The use of hypervisors is one of the present typical solutions for the deployment of VNFs. (Other solution to realize VNFs may include software running on top of non-virtualised software by means of an operating system, e.g. when hypervisor support is not available or VNFs implemented as an application that can run on bare metal.)

### 7.2.3 VNF

VNF is a software package that implements a network function with well-defined functional behaviour and external interfaces. A VNF can be decomposed into smaller functional modules for scalability, reusability and faster response or multiple VNFs can be composed together to reduce management and VNF Forwarding Graph.

#### **VNF Forwarding Graphs**

An end-to-end network service can be defined as a forwarding graph of network functions and end points/terminals.

NFV Forwarding Graphs benefits:

- Efficiency - Compute resources assigned to function and network capacity sized to current load and shareable across functions.
- Resiliency - In some cases, backup function and network capacity can be shared
- Agility - Shorter deployment intervals for upgrades and new features since functions are software based
- Expressiveness - Virtualised switching functions and/or configuration of VNFs can implement forwarding graphs in a more straightforward and efficient manner.
- Flexibility - Reduce configuration complexity. Support new service and business models: deployments in other operator's network, third-party datacentres

The VNF can be composed of multiple internal components, e.g. one VNF can be deployed over multiple virtual machines (VMs), where each VM hosts a single component of the VNF. However, in other case, the whole VNF can be deployed in a single VM.

A large scale VNF may be composed of one or more constituent VNFs connected together in some form of a forwarding graph. An individual constituent VNF can have the following deployment cases:

- 1:1 implementation of single Network Element's Network Function by a single VNF;
- N:1 case where there are N parallel constituent VNFs implementing the capacity of a single Network Element's Network Function;
- 1:N case where N Network Elements Network Functions are implemented by a single VNF (for example, a virtualised home gateway).

Figure 7.6 shows the 1:1 mapping between a Network Element and VNF. In this case then external interfaces and management will likely be completely specified by the existing Network Element standardized interfaces.

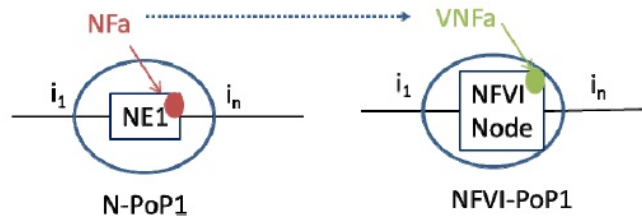


Figure 7.6 - Mapping NE:VNF 1:1 [65]

In this case Network Functions NFa is implemented in NE1 deployed at N-PoP1, and supports interfaces  $i_1$  to  $i_n$ . In this example NE1 provides no network functions other than NFa. The equivalent Virtualised Network Function (VNFa) executes in the NFVI Node and provides the same interfaces  $i_1$  to  $i_n$ . If the NFVI Node implements no other VNFs, then it could be considered a direct replacement for NE1. For a direct replacement of an NE by an NFVI node running the equivalent VNFs at the same location, NFVI-PoP1 would be the same location as N-PoP1.

Figure 7.7 illustrates the case of the N:1 mapping of the Nfa implemented in a single high capacity Network element (NE1) into three instances of VNFa. The instances of VNFa, in general, may be executing in different NFVI Nodes in different NFVI-PoPs, and in this figure they are shown executing in two different NFVI Nodes in two different NFVI-PoPs. In aggregate the combination of NFVI Nodes 1 and 2 and the split/merge functions prove the equivalent set of external interfaces  $i_1$  to  $i_n$ . In this example the splitting and merging functions are allocating traffic across instances of the same type of VNF (VNFa). This sort of splitting and merging function is typically referred to as load balancing.

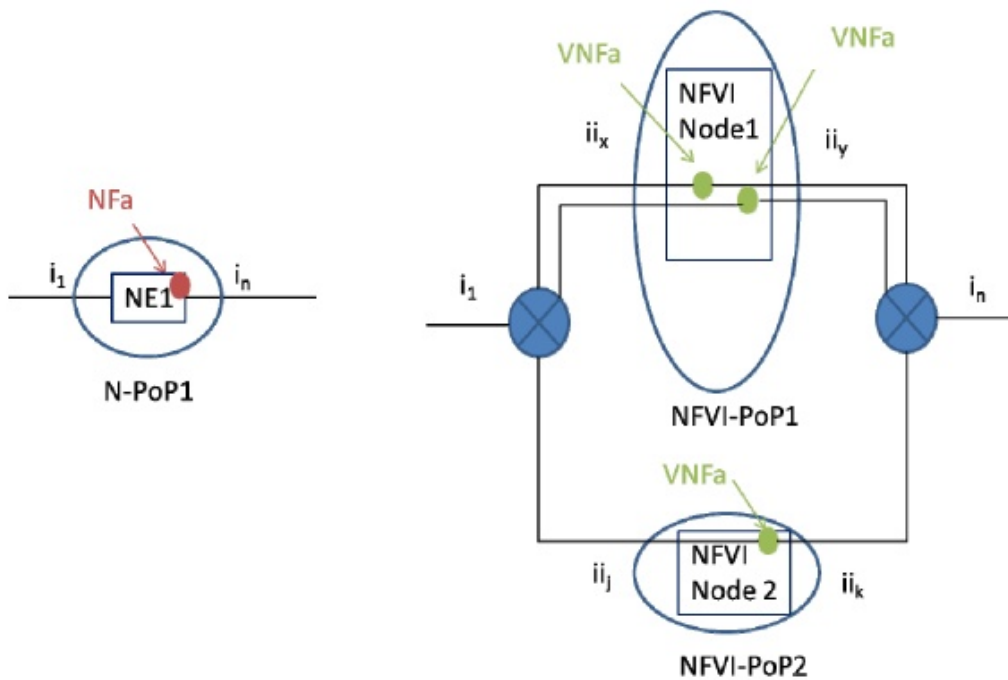


Figure 7.7 - N:1 Mapping of VNF (NFVI-Node):NE [65]

The 1:N case is where N Network Functions are implemented by a single VNF (for example, a virtualised home gateway) and where each individual NF is defined by at least is individual state. Figure 7.8 provides an example of three identical NFs implemented in different NEs in different locations. The equivalent VNF (VNFa) supports a larger number of interfaces ( $i_j$  to  $i_k$ ) from a single instance. The single VNF will operate some means of partitioning between the individual NFs, but in this scenario, this level of partitioning would normally be vendor implementation specific.



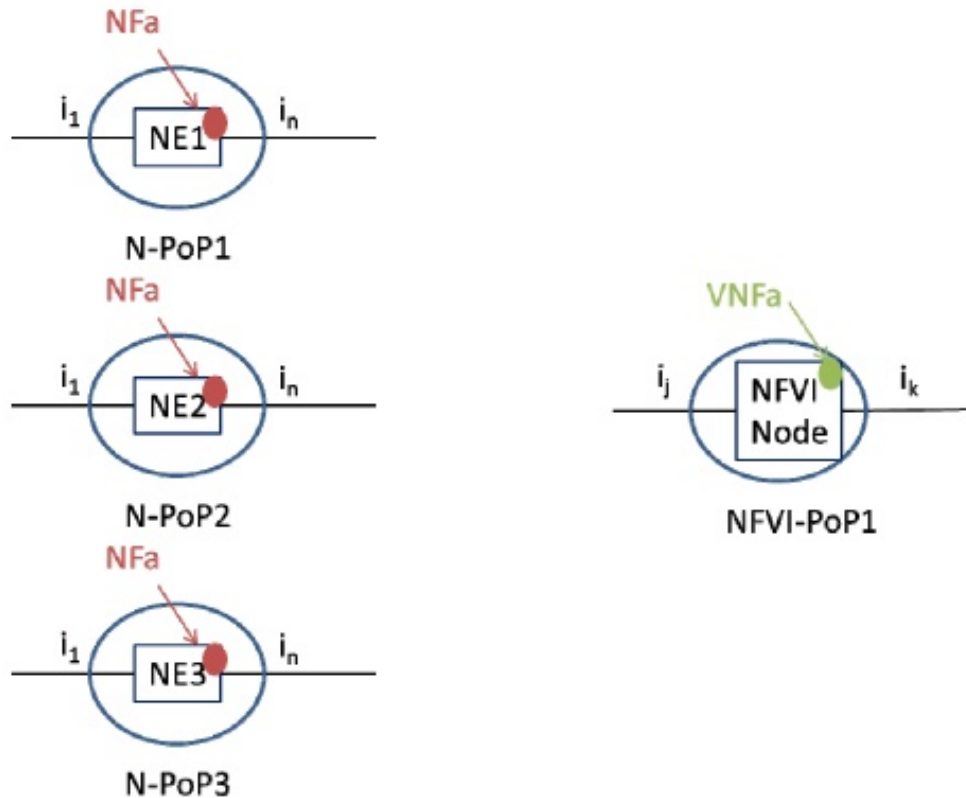


Figure 7.8 - Example 1:N Mapping of VNFs:NEs [65]

Examples of network functions that can be virtualised as VNFs are 3GPP *Evolved Packet Core (EPC)* network elements, e.g. *Mobility Management Entity (MME)*, *Serving Gateway (SGW)*, *Packet Data Network Gateway (PGW)*, elements in home network, e.g. *Residential Gateway (RGW)* and conventional network functions, e.g. *Dynamic Host Configuration Protocol (DHCP)* servers, firewalls, etc. Examples can be found in [66].

#### 7.2.4 NFV MANO (Management and Orchestration)

Management and co-ordination (orchestration) includes three components [67]:

- *NFV Orchestrator (NFVO)* - responsible for coordinating (managing) NFVI resources for multiple *VIMs (Virtualised Infrastructure Manager)*, performing resource coordination functions, network service life cycle management (e.g. management strategy, expansion, feature measurement, event correlation) coordination for network services, overall resource management, endorsement and approval of NFVI applications.
- *VNF Manager (NFVM)* - responsible for managing the life cycle of VNF instances (can be assigned to manage one instance and can also handle multiple instances of the same or different type), overall coordination and configuration adaptations, and event reporting between NFVI and *EMS/NMS (Element Management System/Network Management System)*.
- *Virtualised Infrastructure Manager (VIM)* - responsible for control and management of NFVI compute, storage and network resources in the subdomain infrastructure of a single

operator, collection of performance and fault information of hardware resources, software resources and virtualised resources.

## 8. New generation of multimedia services/applications

### 8.1 Introduction

The environment of the new generation of ICT network platforms provides the wide spectrum of new services and applications. In addition to Internet of things services which are presented in different module this material focuses on following multimedia services:

- Internet multimedia services and applications
- Hybrid Broadband Broadcast TV services
- eServices and mServices
- IoT services and applications
- NGN services
- WebRTC

### 8.2 Internet multimedia services and applications

#### 8.2.1 Software as a Service

*Software as a service (SaaS)* is sometimes presented as a “software on-demand” service. SaaS delivers and licenses software to users. Licensing process is realized on a subscription basis.

Licensed software is centrally hosted and provided to customers over a network, typically the Internet. Interested users use a web browser as a thin client to access this service [68]. SaaS is most used for following purposes: business applications (including office and messaging software), management software, games, computer-aided design software, payroll processing software. In case of conventionally sold traditional software users receive a license which is valid entire their life. Users pay a software price in advance plus some optional ongoing support fee.

The SaaS model includes following benefits:

- generally global accessibility,
- administration is easier,
- compatibility (same version of software at all users),
- collaboration is easier (same version of software at all users),
- automatic updates and patch management.

Few of the most popular online services today are **online picture editors and video editors** (studios). These are typically packed with social media service providing ability to quickly and conveniently share user’s creations on the Internet. Some examples are: YouTube Video Editor, WeVideo, PowToon, Wideo, Weavly, Kaltura, MIXMOOV, Shotclip, Magisto.

**Personal cloud** is a platform where various digital content and information services are concentrated and accessible from any device in Internet. For user this platform (cloud) doesn't appear as a tangible entity. Personal cloud provides users with the ability to upload, store, synchronize, stream, retrieve and share content.

### 8.2.2 VoD streaming

*Video on demand (VoD)* are systems as well as services which allow users to find, select and then watch or listen to favourite audio or video content. This content is available to users anytime, i.e. when they choose and they don't have to adapt and watch at a specific broadcast time.

Users can use personal computers or TV sets to receive video on demand when IPTV technology is used. This is often used scenario. In case of television VoD systems VoD content is streamed directly through a set-top box, PC or other device which allows viewing in real time. VoD content can be also downloaded to a VoD compatible device (e.g. computer). There are several modes of VoD distributions [69]:

- *Transactional VoD (TVoD)* - customers pay for each particular piece of VoD content (e.g. iTunes, Google Play).
- *Subscription VoD (SVoD)* - users are charged a monthly fee to access unlimited programs (Amazon Video, Hulu Plus, Netflix).
- *Near VoD (NVoD)* - TV program is broadcasted in multiple copies at short time intervals (typically 10-20 minutes). Viewers don't have to accommodate for regular scheduled time of program broadcast.
- *Advertising (or Ad-based) VoD (AVoD)* - users don't pay for the content in return for spending time watching ads (e.g. YouTube).

### 8.2.3 Live streaming

Live streaming is a process when multimedia is delivered to a client (user) live over the Internet. Streaming means that multimedia is constantly received by the end user device and then shown to that end user. Streaming is similar to downloading, i.e. it is a process of delivering media but this delivering has to meet special regular conditions. In case of downloading data are available after last byte is received. In case of streaming data (e.g. movie) can be processed (e.g. played by a user's media player) before the entire file has been transmitted. The streaming process must be allowed by convenient audio (MP3, Vorbis or AAC) and video (H.264 or VP8) codec in case of multimedia.

### 8.2.4 Cloud gaming or Game as a Service

Cloud gaming (or Game as a service or Gaming on demand) belongs to online gaming. Currently, we can distinguish two main types of cloud gaming:

- cloud gaming based on video streaming - the games are streamed to user's computers, terminals and mobile devices like video using a thin client (similar to the VoD service),

- cloud gaming based on file streaming - device runs actual game. At the beginning small part of the game is downloaded to user's device and quickly executed so the user can start to play. The rest of the game data is downloaded to the device during playing.

### 8.3 Hybrid broadband broadcast TV services

*Hybrid Broadcast Broadband Television (HbbTV)* represents a consortium of industry companies engaged in digital broadcasting, Internet domain and standardization. HbbTV is also an international standard (specification) defining a delivery of digital interactive TV to the users through a common user interface on TVs or set-top-boxes.

Digital TV can be delivered via broadcast technologies (DVB over cable, satellite or terrestrially) as well as broadband technologies allowing access to Internet. HbbTV is not only about digital TV but brings users a lot of information and entertainment services to augment user experiences. HbbTV tries to combine the best of Television and the Internet. The broadcast connection is mainly used to transmit standard TV, radio and data services (linear content), transport and signalling of broadcast-related applications and associated data and synchronization of TV/radio/data services and applications. The broadband connection is used to carry on demand related content (e.g. VoD), transport of applications and associated data which are or not related to broadcast content (e.g. teletext), serve as a duplex channel for an exchange of information between applications and application servers and to discover broadcast-independent applications.

Currently, smart TV technology offers users digital television and a lot of interactive services. Users can watch linear broadcast (TV or audio) programmes (left part of Fig. 8.1) and they can also activate smart platform (e.g. Samsung Smart Hub) offering access to a number of attractive applications utilizing broadband connection of TV set to provide necessary information (right part of Fig. 8.1 - Portal). However, these applications are called as broadcast independent applications, i.e. they have no close relation to broadcast linear service (content).

Figure 8.1 shows how HbbTV can integrate these applications and some of them to tie up to broadcasting services [70].

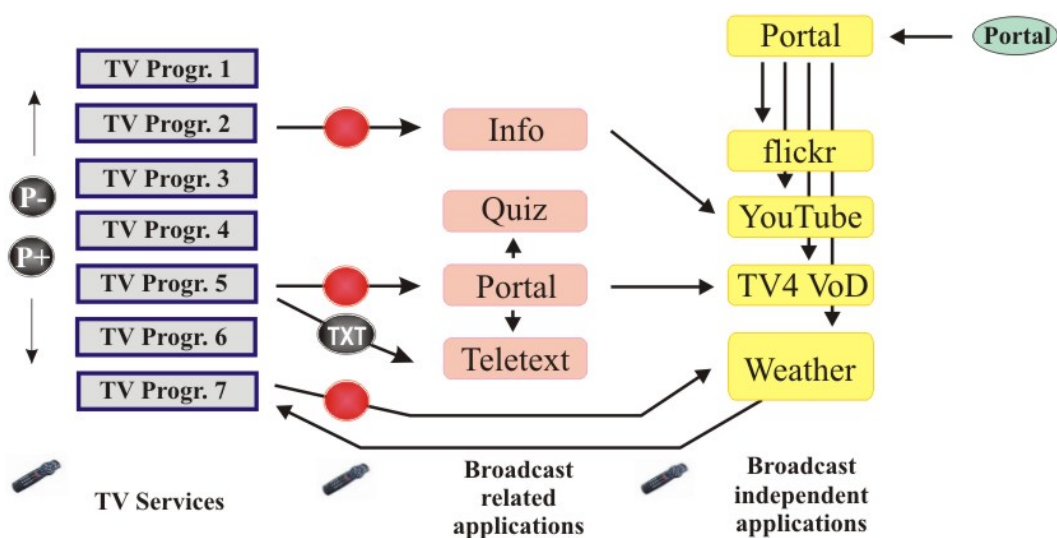


Figure 8.1 - HbbTV service concept [70]

### 8.3.1 HbbTV services

To provide HbbTV services the end user device has to start particular application to allow the users to access all service functions. HbbTV technology extends functions of DVB and smart TV technology by following services [71]: VoD (Catch-Up TV, Start-Over, Push VoD, Live streaming), information services (news, weather, traffic, sport, eGovernment, enhanced teletext, guides, EPG), enhanced TV (additional information on TV programs), games, courses and education, interactive advertising, voting and polling, social networking, home shopping, TV portals, 2nd screen, social and accessibility services, **PVR** (*personal video recording*), personalization.

### 8.3.2 Video (content) on demand services

HbbTV video (content) on demand service is a very attractive service for users because it sets users free from scheduled broadcasting of programs. VoD applications offer users a list with a number of movies, programs, shows and etc. organized and presented in attractive form (GUI).

Example of VoD service is a **push VoD** that provides users with video content on demand. The push VoD system is based on a presence of user's local storage that is usually implemented inside a set-top-box. Selected content is downloaded to this local storage and available to users anytime to watch without a need to wait for buffering. The push VoD system utilizes a *personal video recorder* (**PVR**) service to store selected content that is often transmitted overnight (low traffic) or all day long at low bandwidth.

The push VoD service is suitable for broadcasters and users who lack network connectivity or broadcasters who want to optimize their video streaming network infrastructure because the most popular content is preloaded to the consumer device.

Another application of VoD service is in **catch-up TV** applications. The HbbTV catch-up TV brings users new features of freedom against watching linear TV content.

Users can watch TV channel content no matter it was already broadcasted. The catch-up TV service provides users with access to an archive with television shows and other TV content for a certain period of days after their original television broadcast.



Figure 8.2 - ARTE channel with catch-up features

A **start-over** service is another HbbTV service which should be also very interesting and valuable function for end users. It is also called as a Restart function. This function comes in handy mainly when you came to watch a favourite program that already started broadcasting some time ago. Using this function you simply restart broadcasting of a program and you don't miss its beginning anymore.

However, this function is only active for given program during its broadcasting (from e.g. second minute until the program is finished). When this program is over the user can use catch-up service to watch it from an archive. The start-over function can be limited to certain period of day. Users can also return back to the live session.

### 8.3.3 Other HbbTV multimedia services

Information services provided by HbbTV applications are equipped by an attractive GUI and allow users to browse various theme oriented information (news, weather, exchange rates, stock market, sports, traffic, eGovernment). Thanks to HTML this GUI can shows texts, pictures, graphs, maps and even videos. Similarly, the electronic program guide can be enhanced by various video shots.

HbbTV v2 defined support for *companion screens* (CS) applications. Using HbbTV application on TV set users can launch CS application on other device. These applications can communicate each other. There is also possibility for an application on companion screen to discover HbbTV terminal and launch a broadcast independent CS application on it.

## 8.4 eServices and mServices

With emerging the *information and communication technologies* (ICT, especially Internet and web technologies in the last decade) new type of services started to appear. These services are called electronic services (e-services). One of several definitions of e-services claims [72]:

- an e-service is any asset that is made available via the Internet to drive new revenue streams or create new efficiencies.

E-services distinguish three main components: service provider (public agencies, universities, commercial companies, etc.), service receiver (citizens, students, firms, etc.) and delivery channel (i.e. technology used - Internet, television, telephone, radio, CD-ROMs). E-services can help in accessing broader customer base. They can be available 24 hours a day and accessible from anywhere. Installation and operation costs can be significantly decreased.

**E-commerce** or **electronic commerce** is a service covering online business activities related to products and services. Very simply said it is about selling and buying the goods realized over Internet, i.e. particular parties interact mostly in electronic way (than by direct physical exchanges). It also covers every business transaction conducted over ICT and resulting in transfer of ownerships and copyrights for using various goods and services [73].

**E-business** represents a complex application of ICT into all parts and processes of the business world. All e-commerce activities are within e-business extended by internal business processes (inventory management, risk management, production and product development, finance, human resources and knowledge management). Examples of e-commerce services: e-shopping, e-banking, payment systems, digital wallet, automated online assistant, online reservations and electronic tickets, shopping cart software, etc.

E-commerce can make use various sales scenarios:

- *business-to-business (B2B)* - companies sell their products (services) online to other companies.
- *business-to-consumer (B2C)* - companies sell their products and services online to end consumers (anonymous clients). Examples are Amazon, Zappos, MALL.
- *consumer-to-business (C2B)* - customers (consumers) offer their products or services online to companies.
- *consumer-to-consumer (C2C)* - consumers (people, citizen) offer and sell online their goods directly to other consumers (people). Examples of C2C are eBay, Amazon, BrickLink which use a PayPal system.
- *government-to-business (G2B), business-to-government (B2G), government-to-citizen (G2C), citizen-to-government (C2G)*, etc. - other e-commerce scenarios where transactions are realized with the government.

**E-government** is a term which includes usage of various tools, methods and information and communication technologies with aim to provide and improve public services for enterprises, companies and citizens [73].

It is e-business service in public sector. E-government delivers government services to companies (G2B), citizens (G2C), employees of public administration (**G2E**, *government-to-employee*) and among various government organizations, institutes and departments (**G2G**, *government-to-government*).

The main objective is to bring the public administration closer to citizens and companies in an efficient and cost effective manner. Its main idea is to provide citizens with constant access to public services as well as to improve efficiency of (internal) public administration operation.

**E-signature** is the electronic version of a handwritten signature which is associated with a person indicating his adoption of the document.



It can be a digitized image of a handwritten signature, a symbol, voiceprint, etc., used for identification of authors of an electronic document, message or report. E-signature is vulnerable to copying and tampering, and needs proprietary verification software. On other hand, a **digital signature** relies on a special mathematical scheme which ensures authenticity of the document.

**Electronic banking** allows customers to access bank from anywhere (and to maintain permanent access) and they can acquire lower transaction costs.

**eHealth** can be defined as the usage of modern information and communication technologies to provide correct healthcare information in right time at right place to improve healthcare process and quality of life and to meet the needs of citizens and patients, healthcare professionals and providers, and policy makers [74].

It is based on digital data (patient records) which are electronically transmitted, stored and retrieved for clinical, educational and administrative purposes. eHealth covers for example: communication of patient (health) records between healthcare professionals, e-consulting, ePrescribing (access and printing patient prescriptions), diagnostic tests, diagnosis, treatments and telemonitoring at a distance, information services, mHealth, healthcare management system.

**Electronic learning (e-learning)** can be characterized as an application of ICT into development, distribution and management of educational process.

E-learning covers various forms of education such as web education, distant education, e-teaching, computer supported education, virtual classes, m-learning, cooperation. The educational process is usually realized via Internet, Intranet, audio or video conferences, terrestrial or satellite broadcasting, media such as CD or DVD ROMs, USB flash drives. E-learning also represents a form of self-education through electronic training materials distributed by aforementioned channels. It can be also a part of combined form of education. The educational process is often delivered, tracked and managed by *learning management system (LMS, e.g. Moodle)*.

**Teleworking, telework, telecommuting, e-working** or remote/distance working represent a form of working when worker doesn't have to commute to a central place of work.

Although, a lot of workers work from home some workers can work at various places (shops, abroad). Current teleworker utilizes computer for a work which is connected to the company network.

Teleworking reduces operation costs and enhances the labour productivity and results.

However, teleworking has also some drawbacks. It puts higher stress on worker's motivation to work. Distractions at home can be finally more critical than in work (e.g. children, pets, and neighbours). Teleworker can loose a professional contact with nonteleworkers.

## **8.5 Internet of things applications and services**

**IoT** (*Internet of things*) is an emerging global Internet-based technical architecture facilitating the exchange of goods and services in global supply chain networks has an impact on the security and privacy of the involved stakeholders [75].

The number of applications and services that can provide IoT is practically unlimited and can be adapted to many fields of human activity by facilitating and enhancing their quality of life in multiple ways.

Some of IoT applications and services are as follows:

- **Connected intelligent buildings:** Improvements in efficiency (energy management and saving) and security (sensors and alarms). Domotic applications including smart sensors and actuators to control home appliances. Health and education services at home. Remote control of treatments for patients. Cable/satellite services. Energy storage/generation systems. Smart thermostats. Smoke detectors and alarms. Smart door locks. Safety for all family members.
- **Smart cities and transportation:** Integration of security services. Optimization of public and private transportation. Parking Sensors. Smart management of parking services and traffic in real time. Smart management of traffic lights depending on traffic queues. Locate cars that have overstayed Smart energy grids. Security (cameras, smart sensors, information to citizens). Water management. Parks and Gardens irrigation. Pollution and mobility controls. Get immediate feedback and opinions from citizens. Accident monitoring, emergency actions coordination.
- **Education:** Linking virtual and physical classrooms to make learning more efficient and accessible, e-learning. Access services to virtual libraries and educational portals. Interchange of reports and results in real time. Lifelong learning. Foreign languages learning.
- **Consumer electronics:** Smart phones. Smart TV. Laptops, computers and tablets. Smart refrigerators, washers and dryers. Smart home theatre systems. Smart appliances. Personalization of the user experience. Personal locators. Smart glasses.
- **Health:** Monitoring of chronic diseases. Improvement of the quality of care and quality of life for patients. Activity Trackers. Remote diagnostic. Connected bracelets. Interactive belts. Sport and fitness monitoring. Drug usage tracking. Biochips. Brain-computer interfaces.
- **Automotive:** Smart Cars. Traffic control. Advance information about what is broken. Smart energy management and control. Self-diagnosis. Accelerometers. Position, presence and proximity sensors. GPS tracking. Vehicle speed control.
- **Agriculture and environment:** Measurement and monitoring of environmental pollution (CO<sub>2</sub>, noise, contaminant elements presents in ambient). Forecasting climate changes based on smart sensors monitoring. Sensors in pallets of products. Waste management.
- **Energy services:** accurate data on energy consumption. Smart metering. Smart grids. Analysis and prediction of energy consumption behaviours and patterns. Forecasting future energy trends.
- **Smart Connectivity:** Data management and service provisioning. Access to email, voice and video services. Interactive group communication. Real time streaming. Interactive gaming. Augmented reality. Network security monitoring. Biometric authentication methods. Consumer telematics. M2M communication services. Virtual reality. Computer vision.

- Manufacturing: Gas and flow sensors. Smart sensors of humidity, temperature, motion, force, load, leaks/levels. Machine vision. Acoustic and vibration sensing. Compound applications. Smart control of robots. Pattern recognition. Machine Learning.
- Shopping: Intelligent shopping. Inventory control. Control of geographical origin of food and products. Control food quality and safety.

## 8.6 NGN services

### 8.6.1 VoIP

*Voice over IP (VoIP; or IP telephony, Internet telephony)* is a set of technologies necessary for voice communications delivering and multimedia sessions providing over **IP** (*Internet Protocol*) networks (Internet).

The Internet telephony represents the delivery of communications services such as voice, fax, SMS, voice-messaging over Internet, rather than via the *public switched telephone network (PSTN)*. VoIP telephone call setup process is similar to traditional digital telephony and includes actions: signalling exchange, channel setup, analogue voice signals digitization, voice data encoding. The encoded voice data are packetized and transmitted as IP packets over a *packet-switched data network (PSDN)*.

Examples of VoIP applications are Skype, Google Talk.

There are several competing approaches how to implement the VoIP. Each one is based on a set of protocols to handle signalling, data transmission, and other tasks. The most used protocol in a VoIP world is SIP [76]. The *Session Initiation Protocol (SIP)* is a communication protocol which provides signalization of control for multimedia communication sessions. SIP is an application-layer control protocol that handles the setup, modification, and tear-down of multimedia sessions.

### 8.6.2 Hosted Call Centres

Over the last decade contact centres have experienced an extensive evolution overall. Many companies utilize a number of contact centres to manage all interactions with their customers (whether it is an in-house team or outsourced to third-party assistance). Hosted VoIP telephony is quickly becoming the standard communications platform for organizations of all sizes.

The wholesale transition to feature-rich Hosted VoIP service (from traditional telephone systems) has already begun and offers considerable benefits:

- immediate cost savings,
- increase in system reliability and worker productivity.

Deployment of Hosted VoIP technology requires little on-premises equipment. In most cases, needed equipment is limited to a high-quality router, *Integrated Access Devices (IADs)*, and IP telephones. Analog telephones can be also used in some cases, but IP telephones are strongly recommended because they

- offer more functions,

- require less hardware,
- are easier to use.

### 8.6.3 IPTV

ITU-T defines the **IPTV** (*Internet Protocol Television*) by the following definition [77]:

- IPTV are multimedia services such as television/video/audio/text/graphics/data delivered over IP-based networks managed to support the required level of **QoS/QoE** (*Quality of Service/Experience*), security, interactivity and reliability.

In other words, IPTV is a system which delivers (using streaming technique) television services using IP suite over PSDN networks (LAN, Internet) instead of being delivered through traditional terrestrial, satellite or cable television systems [78]. End to end chain for delivery of the IPTV content to the end user usually contains these 4 main domains that are involved in the provision of an IPTV service (Fig. 8.3): Content provider, Service provider, Network provider and End-user.

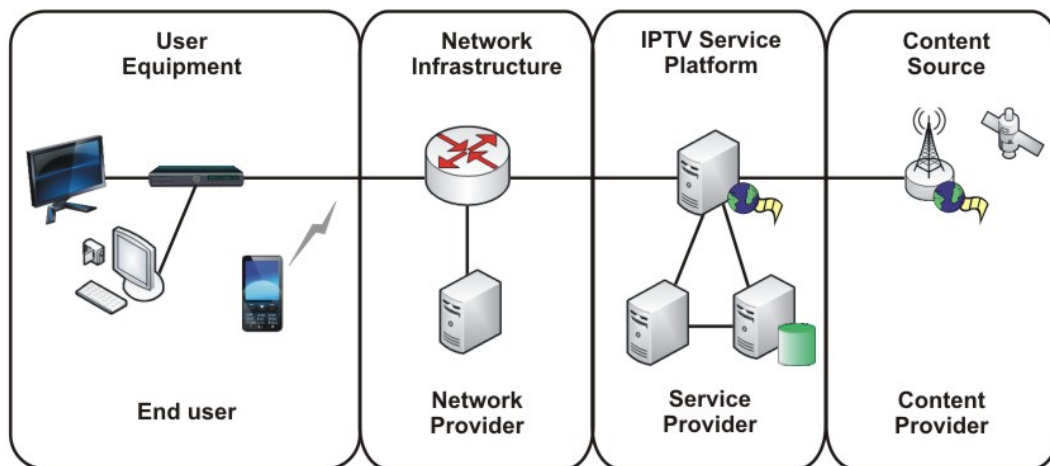


Figure 8.3 - IPTV domains

### 8.6.4 VoIP VPN

Combination of two technologies: the voice over IP and virtual private networks provides **VoIP VPN** technology which offers a delivery of secure voice. As was already mentioned, the VoIP transmits a human voice as a digital data stream.

Then it is quite easy to provide voice encryption via VPN tunnels just by applying standard data-encryption algorithms which are inherently available in protocols used to implement VPN tunnels. Application of the Voice over IP through VPN brings another benefit, however. It is difficult to pass the SIP protocol through a firewall because it uses random port numbers for connections setup. A VPN is good solution how to avoid firewall problems when configuring remote VoIP clients.

## 8.6.5 Supported service (ISDN services emulation/simulation)

During evolution to NGN, NGN shall support legacy terminal equipments (e.g. PSTN/ISDN phones) and PSTN/ISDN-like capabilities.

PSTN/ISDN Emulation:

- from the end user perspective, the NGN “appears” supporting the same types of services offered by the existing PSTN/ISDN,
- legacy terminals can use existing telecommunication services while connected to NGN.

PSTN/ISDN Simulation:

- NGN terminals in an NGN network can use PSTN/ISDN-like service capabilities,
- legacy terminals with terminal adaptations may be used too,
- Implemented over IP-based control infrastructure (e.g. using SIP).

## 8.7 WebRTC

*Web Real-Time Communications (WebRTC)* is a collection of open standards for real-time communication, mainly developed by the *WebRTC World Wide Web Consortium (W3C)* Working Group and the *Real-Time Communication in Web-browsers (RTCWEB)* *Internet Engineering Task Force (IETF)* Working Group.

The W3C focuses their work with WebRTC mainly on the browser *Application Programming Interfaces (APIs)* to interact with the audio/video sources. The IETF created the RTCweb group to focus on the inter-browser interface and definition of (signaling) protocols. WebRTC [79] opens up possibilities for real-time communications such as audio and video calls, screen sharing and video conferencing within web-browsers, but without any use of additional software (only modern web browsers are required). This makes it easy for web developers to implement WebRTC features only by using *Hypertext Markup Language version 5 (HTML5)* and a variety of JavaScript *Application Programming Interfaces (API)*.

Besides providing a powerful decentralized media engine within browsers, WebRTC has other benefits such as open-source code & APIs, free audio & video codecs (adaptive, high definition), and network support inbuilt (e.g. encryption, network discovery). However, due to its technical design, WebRTC is not narrowed to the use within browsers. It can also be used via apps and native implementations, so that nearly every modern connected device - computers, tablets or even televisions - could become a WebRTC peer and thus a fully-fledged communications device. In contrast to other real-time communication systems, providing communication using WebRTC does not require a big infrastructure which handles the communication traffic throughout the peers.

### 8.7.1 Applications

With WebRTC [80], various applications can be built - not only limited to communication features. WebRTC is not (only/mainly) about “calling” from within the browser, but about enabling web developers to access to audio/video input devices via JavaScript as well as abstracting the problem of browser-to-browser communication for ordinary web developers.

Once the browser-to-browser communication problem has been solved, WebRTC provides both a user data channel for real-time communications data, but also a data channel to send any kind of other data in a peer-to-peer manner.

All of this does mostly not require plug-ins - but is natively supported in the browsers (currently Google Chrome, Mozilla Firefox, Opera, Microsoft Edge).

### **Browser-to-browser applications for voice calling & video chat**

The simplest application for WebRTC is the audio/video communication between browsers. The inbuilt WebRTC capability provides microphone (audio) and camera (video) access (the user can select the device and grant permission).

The important API functions for this use case are: `MediaStream/getUserMedia` (HTML 5) and `RTCPeerConnection`. The `getUserMedia` API adds access to dynamic sources such as microphones and cameras. The characteristics of these sources can change in response to application needs. The `PeerConnection` is a media technology that allows two users to communicate directly, browser to browser. This communication is coordinated via a signalling channel which is provided by unspecified means, but generally by a script in the web page that has been provided by the web server. Sample services are: [appear.in](http://appear.in), [talky.io](http://talky.io).

### **P2P file sharing**

The `RTCDataChannel` lets a web application send and receive generic application data peer-to-peer. The `DataChannel` interface represents a bidirectional data channel between two peers. While the `PeerConnection` is a channel for RTC only, the `DataChannel` can transport any type of data. A sample service is [sharefest.me](http://sharefest.me).

### **Screen sharing**

The `getUserMedia` API can not only access camera/microphone as media source, but also the shared screen. For security reasons, accessing the screen requires a plug-in. Most services that are used to communicate with audio/video also offer screen sharing.

### **Collaborative whiteboard**

Besides A/V communication and screen sharing, the applied data channel can also be used to transfer not only files, but also control information. This control information can be used to modify displayed browser content. A sample application for this can be a collaborative whiteboard. By sending the inputs from one whiteboard (“editor”) to all other whiteboards under the same link (“viewers”) the browser application can act as shared whiteboard.

### **Conferencing**

From the pure browser concept, WebRTC is conceived as peer-to-peer communication, without requiring additional infrastructure. This architectural approach makes it difficult to realize sessions with multiple streams such as group video conferences or other "n-to-m" broadcasting scenarios. This is the place, in which the **conferencing** building block comes into play. The conferencing building block cares about the distribution of media traffic to a group of peers. That distribution is possible in three different ways, which are primarily differing in their requirements of additional servers.

## 9. Information and network security

In general, the security can be defined as the quality or state of something's existence (e.g. system, information) when there are no dangers or threats.

In order to achieve this aim (state) it is necessary to apply various security measures. In case of a security of the **ICT** (*information and communication technology*) systems it is about deployment of such measures and strategies which provide the proper operation of those systems and a protection of everything they contain, i.e. everything what has certain value (e.g. information, hardware). The ICT systems security will include various security areas [81]:

- information (data) security - protecting information (data) from eavesdropping and misuse, destruction and modification,
- computer security - secured operation of computers and a protection of data which are processed and stored by computers (end devices),
- network security (access, communication and system) - a protection of data during their transmission via communication environment and a protection of computers connected to computer networks.

It is possible to encounter other security areas such as physical security (protection of physical objects), personnel security (who is authorized to access and use secured objects and items), software security including application security (protection of operating systems, applications and various tools deployed for provision of the information security) or Internet security. These areas can be partially or completely covered by aforementioned types of the security.

Currently, we very often meet a term cyber security. It is the security of a cyber space - a virtual space composed of worldwide interconnected communication networks, information systems and various other subsystems (control, manufacture, security, etc.) including hardware, software and data (i.e. it is built on the real space = Internet) where people, software and services mutually interact.

### 9.1 Terminology in information security

In this section we define basic terms which are used in the information security area and which are also used in standards and recommendations defined for ICT area.

If we want to implement a security in ICT (in general) we have to realize and identify **what** we want to protect. It means we have to define everything what has a value for us (as persons) or for organization (company) and all of that will be jointly referred to as **asset/assets**. In order to achieve a required level of asset protection the company or organization must specify a set of security rules. Strategy how to achieve this security level is described by a **security policy** which represents a set of recommendations, rules and methods which define how to manage, protect and propagate sensitive information and other assets. The compromise or misuse of the **sensitive information** can cause significant loss and damage to persons, organization or company.

The company assets include:

- **tangible assets** (hardware, buildings, etc.),

- *intangible assets*,
  - information (data),
  - software,
  - ability to provide a product, service, information,
- *personnel* (operating information system).

**Vulnerability** represents a fault or a weakness of the information system that can be exploited by threat in order to cause a loss or damage of the information system.

It is important to note that the vulnerability itself doesn't cause a system damage it only creates conditions for realization of threats on the information system. System weaknesses can exist at all levels of the information system (physical layer, hardware, software, organizational or personnel structure/policy, information system management). Examples are: vulnerabilities of operating systems, weak access passwords, software fault, unsecured system ports, unlocked doors, etc.

Danger that jeopardizes assets is called a **threat**. Threat represents possibility for exploitation of information system vulnerability in order to invoke a system damage or failure. This damage can happen because of attacks on information, a system alone or other system sources which can cause a system failure.

If system vulnerability is exploited for the purpose of causing damage to the system assets we speak about an **attack**. We distinguish intentional exploitation of vulnerability (targeted attack causing damage) and unintentional action which causes asset damage (e.g. flash can make damage to electronic devices or buildings).

The **risk** represents a probability that something undesirable happens in a system. It represents a potential possibility that some kind of a loss (damage) occurs in the information system because a threat exploits the vulnerability of some system component.

At present time, there exist a several important asset (information, data, software, hardware, services) attributes/properties which should be protected by the ICT security of their violation.

First one of the most important information security models has defined three fundamental components (requirements) [81]:

- **confidentiality** - assets are only available to authorized subjects,
- **integrity** - assets can be only modified by authorized subjects,
- **availability** - assets are available for authorized persons when they are required.

Over time this model has been extended by other specific components which should be also taken into account, such as [81], [82]:

- **authenticity** - asset origin is verified, assets are not copies (falsification),
- **accountability** - ability to monitor, what and when the user, system or process performed,



- **non-repudiation** - no party can disclaim later that e.g. given document already signed (digital signature),
- **possession** - quality or state of ownership or control over information,
- **utility** - information is in a state (format) useful for a user,
- **privacy** - secrecy of personal information (credentials),
- **reliability** - device/system behaves consistently.

From general point of view security attacks can be divided into following groups:

- attacks on availability - energy supply interruption, information system overloading,
- attacks on confidentiality - eavesdropping, information flow analysis, copying the data memory, data or device theft,
- attacks on integrity - modification of software, viruses, Trojan horses, logic bombs, modification of stored data, modification of data during their transmission,
- attacks on authenticity - **MITM** (*man in the middle*) attacks, replay attacks, inserting a false record to a database.

It is necessary to realize that there is no absolutely secured information system and security policies attempt to decrease a probability of successful attacks on the information system.

## 9.2 Security mechanisms

Security services are implemented via security mechanisms which support and realize specific operations targeted to protect assets from attacks or their consequences. Some important security mechanisms (based on X.800 recommendation) which are directly implemented at some layer of the protocol (network) model are [83]:

- **encipherment** - concealment of message content; Specific cryptographic transformation is used which converts a message into a form not readable by unauthorized subject.
- **digital signature** - application of cryptographic transformations for provision of message source authentication and data integrity.
- **access control** - the control and verification of access rights to system resources and services is provided based on authentication or another entity information.
- **data integrity** - control mechanisms detect whether transmitted data has been modified during transmission.
- **authentication exchange** - process for exchange of the authentication information between a user (entity) and the information system. It is used for user identity verification and its result influences access control mechanisms.
- **traffic padding** - gaps between transmitted data are padded by additional bits to make a data flow analysis impossible.

- **routing control** - selection of secured (physical) transmission routes for specific data with possible routing change if security violation is expected.
- **notarization** - third party is used for protection of certain aspect of data exchange in the information systems.

Except for aforementioned mechanisms there are also other mechanisms which are not related to a concrete layer of the network model and they can be considered as an aspect of security management. Examples are: security label, security events detection, security recovery, etc.

### 9.3 Security attacks

Security attacks can be divided into two main categories as follow [84]:

- **passive attacks** - aim is to obtain information from the information system but system resources of the information system are not affected. That's why it is very complicated to discover this type of attacks and it is more efficient to apply prevention (e.g. encryption).
- **active attacks** - aim is not only to obtain information but also modification of data flow (traffic) resulting in affecting the system resources (eventually their operation). Therefore this type of attacks can be easier to discover but protection is very complex and efficient protection is centred on attack detection and elimination of consequences caused by attacks.

**Passive attacks** on the system security (Fig. 9.1) mainly make use of eavesdropping or traffic monitoring and they apply two main approaches:

- **release of message content** - usage of monitoring (observing) functions but they are only efficient in a case when communications run in an open form. Main protection mechanisms for communications (data traversed the network) are based on the data encryption (ciphering).
- **traffic analysis** - sophisticated methods which are centered on information collecting from a captured traffic flow by its analysis. Even when data encryption is applied an attacker can deduce information based on patterns or behaviour of the traffic flow. Information deduced by this type of approach can contain e.g. a frequency and lengths of transmitted messages. An attacker can identify a character/nature of user's communications.

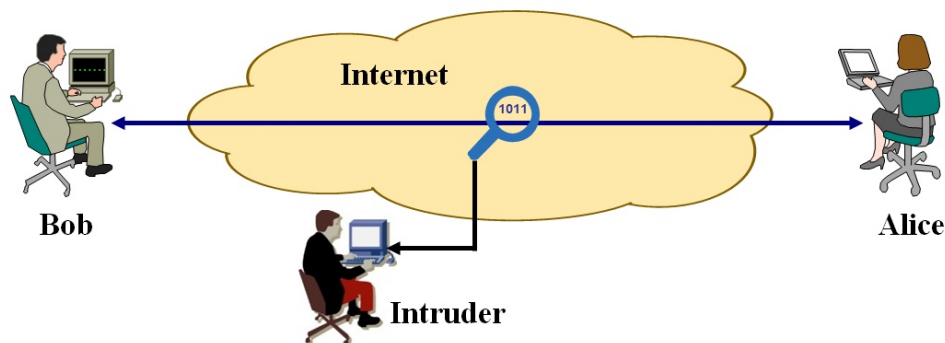


Figure 9.1 - Passive attacks principle

**Active attacks** modify a transmitted data flow or create a new (false) data flow (Fig. 9.2). It is possible to divide them into four categories:

- **masquerade** - one subject pretends an identity of another subject. Authentication information exchange of an authorized subject is captured by an unauthorized subject thereby allowing him to acquire privileges assigned to that authorized subject.
- **replay** - data (packets) from the traffic flow are captured and all of them or part of them are sent to the information system with certain delay causing that system to behave improperly.
- **modification of messages** - original transmitted message is intercepted by an attacker, modified and then resent to a destination. The attacker can obtain message content and by message altering he can even acquire privileges (permissions) of another subject.
- **denial of service** - aim is to prevent from standard service usage or network management. The attacker can cause an unavailability of a service, device (host) or even entire network for other users. An attack to a server realized by multiple computers is more efficient variant of these attacks. It is referred as a **DDoS** (*Distributed Denial of Service*) attack.

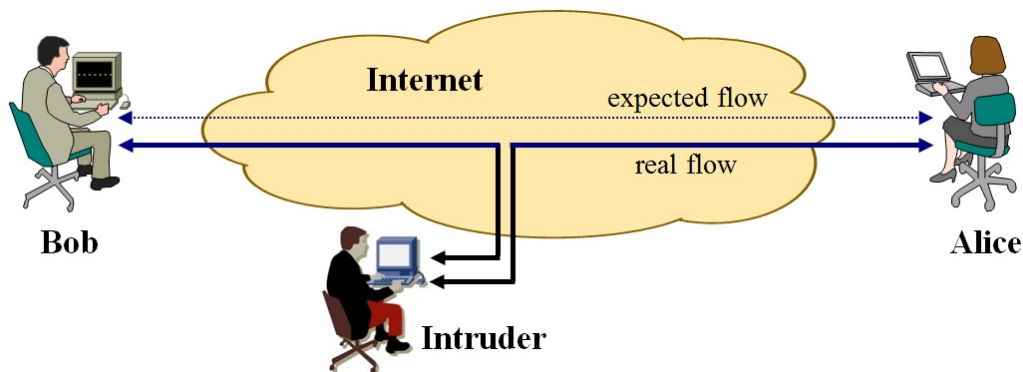


Figure 9.2 - Active attacks principle

## 9.4 Intruders

**Intruder** is a subject/individual which realizes one of the attacks on the information system security. It is a person who obtains or wants to obtain unauthorized access or unauthorized permissions in computer systems.

Currently, intruders don't have to be only computer professionals (specialists) who are able to perform very dangerous attacks with serious consequences for the computer systems and networks but they can be also amateurs with lower level of computer skills (so called wackers) who are able using various tools available on the Internet to perform less dangerous attacks. Based on an attacker position we can distinguish two types of attackers:

- **insider** - usually a subject (person) who is legitimately connected to the internal computer network and who wants to obtain an unauthorized access to data, system resources and services.

- **outsider** - usually a subject (person) who doesn't have an authorized access to the internal computer network and who wants to infiltrate into this network by attacks on its vulnerabilities.

The attackers can act as individuals or as organized groups. In the information security world we can encounter various names for attackers [85], such as:

- **hacker** (white hat) - a person with very high skills in the ICT domain. His activities are mainly oriented to searching for vulnerabilities and security holes of the existing ICT systems. His motivation is excitement, satisfaction and credit. If he informs a system owner about discovered vulnerabilities his activity can be helpful.
- **cracker** (black hat) - a person who is able to breach anti-piracy protections of computer programs. It is a person who makes use of his knowledge in an unethical way. He often concentrates on gaining an unauthorized access to the systems or programs. Based on some sources a cracker is a hacker with malicious intentions.
- **spammer** - a person who sends a large number of unsolicited email messages (viruses can be also used for this purpose).
- **phisher** - a person who tries to obtain sensitive information (e.g. passwords, credit card numbers, etc.) from users using emails or web pages.
- **scriptkiddies** - computer system users with low level of ICT skills who realize their attacks using known scripts containing a code which is able to exploit computer system vulnerabilities. They can cause considerable damage on systems.

## 9.5 Malicious software

**Malicious software (malware)** is a purposefully developed computer program which represents a software threat to the computer system and which can cause serious damage to this system.

This software can be divided into two basic groups based on a distribution mechanism. First group contains the malicious software which needs a host program for own distribution whereas malicious programs in the second group are independent (of any host program). The first group, sometimes called as parasitic, contains

- **viruses** - programs which attach themselves to other programs (executable file) and which can carry out undesirable activities. They are able to attack other files, replicate and distribute themselves and cause damage to other computer systems. They are activated when a host (infected) program is executed.
- **logic bombs** - programs which are secretly and intentionally integrated into standard programs. They are activated when certain conditions are met e.g. an execution of particular application, specified time elapses, at given time, etc. After activation they can show a (false) message, delete or destroy the data, stop running calculations (programs) or perform some other type of malicious operations. It is the oldest type of the malicious software (viruses).
- **Trojan horses** - programs which provide useful functions to users but they contain secret functions as well which perform undesirable and destructive operations in a background. They make use of legitimate privileges of given entity for own activities and they can

e.g. delete users data or gather passwords entered by keyboard (keyloggers), monitor and log user's activity and send this information to an attacker.

- **trap/back doors** - they are Trojan horses which enable an attacker to gain access to the computer system by avoiding the security mechanisms. They can be directly implemented by programmers and used during program debugging and testing phases. They can also be implemented with clear intention to breach the system security (e.g. in games and other useful programs).

Majority of viruses have a life cycle consisted of four phases: dormant phase, propagation phase (virus replicates or makes its identical copies), triggering phase (when specific condition is met) and execution phase (activities often causing damage on the infected computer system). Viruses can be categorized as follows [81], [84]:

- parasitic viruses - viruses distributed using the executable files (file with extensions such as exe, sys, com, bat),
- macro viruses - they most often infect files created by MS Office applications,
- boot sector viruses - viruses which use a boot sector for loading into the system,
- stealth viruses - realize active protection against its disclosure,
- polymorphic viruses - they create modifications (mutations) of the original viruses to make their disclosure more difficult,
- resident viruses - viruses which store themselves within the computer memory,
- encrypted viruses - a part of the virus code is encrypted to make their disclosure more difficult,
- e-mail viruses - they are attached to the emails and activated by opening the attached files.

The second group of viruses which don't need a host program for their distribution are:

- **worms** - are able to replicate and propagate from one computer system to another computer system if these systems are connected to the same computer network. They can propagate as email attachments or they are able to log in to a remote system as a normal user and copy themselves there (and then activate) or eventually they exploit some hole in the system.
- **zombie** - they propagate via computer networks (Internet) and after successful penetration of the computer system they allow to take control of infected system. Several computers infected by the same type of this malware create a botnet. Botnet can be controlled from one remote computer to realize e.g. DDoS attacks.

## 9.6 Network security models

If we talk about the network security two models are the most often specified [84]:

- network data (information) security model,
- network access security model.

In general, it is known that a public network which is used for mutual interconnection of all users in Internet is an unsecured network. If we want to transmit sensitive information via the public network it is necessary to apply the network data security model (Fig. 9.3). This model is based on 4 main requirements:

- design of an effective algorithm for security-related transformation of transmitted message,
- generation of keys (secret information) which are used by the security algorithm,
- development of methods which are used for distribution and sharing the keys (secret information),
- definition of protocols for deployment of required security services.

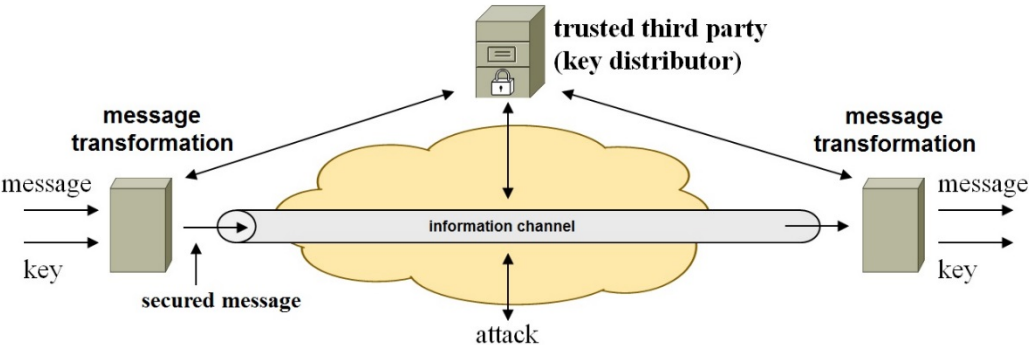


Figure 9.3 - Network data (information) security model principle

The network access security model that is depicted in Fig. 9.4 and that protects the information system from undesirable entries (attacks) requires:

- design of suitable functions for a firewall which will prevent unauthorized subjects and malware from entering a firewall,
- monitoring the communication traversing a firewall (analysis, detection).

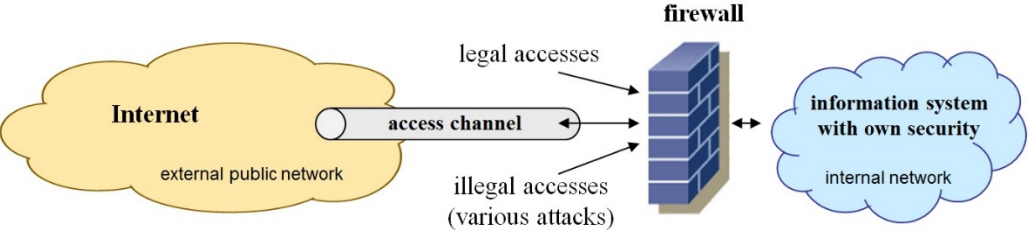


Figure 9.4 - The network access security model principle

The first network security model utilizes various cryptographic algorithms and protocols and can also incorporate a trusted third party in its operation. A core of the second network security model is a firewall. This education material concentrates further on this model and in more details on possible realization of firewalls.

## 9.7 Firewalls

A **firewall** [4] is a network security device that provides the security for an internal private (company/home) network by separating that private network from rest of the world (Internet).

If the company wants to implement the network security it must define its security policy, i.e. a set of rules which provide a protection for its property, computer systems, personal information and other sensitive data (assets). Based on this policy the security functions are defined for the firewall which can be realized by hardware, software or by their combination. The fundamental condition is that all traffic/communication from and to this internal network must traverse the firewall. Then the firewalls can apply two main functions: block or permit communication. Usually, all communication from the internal network to the public network (Internet) is permitted and on the other hand usually almost all requests initiated in Internet that come to the internal network are blocked (dropped). However, in the first case a security policy can e.g. block connections to untrusted sites or other sites which are considered to be a security risk or inappropriate for companies. As was already mentioned the firewall must be also at the same time resistant to attacks directed to itself (it must use strong operating system).

Advantages of firewalls can be summarized as follows:

- they can prevent external users from access to the internal network resources using unsecured protocol by dropping packets destined to its TCP/UDP port,
- they can prevent certain IP addresses from reaching the internal network, e.g. they can forbid WIRTE commands of FTP protocols and only allow to read an FTP content,
- it is more cost effective to use one or few firewalls to secure the private network than to secure every host and its applications,
- it is easier to secure firewalls than end systems. Firewalls have simplified operating systems and these systems do not run complex applications. The less complex an application is, the fewer programming bugs are potentially present in it.

Disadvantages of firewalls can cover:

- the firewall is a central point exposed to attacks. If an attacker takes control of the firewall he can perform attacks directly from the firewall to the internal network as well as the external network or he can block all traffic traversing the firewall.
- firewalls represent network bottlenecks and single points of failure as well. In the case of a firewall failure the entire network loses Internet connection.
- firewalls cannot protect the internal network from internal threats and unfortunately most of attacks originate in the company internal network. In order to protect the internal network from these kinds of attacks it is necessary to deploy other security systems.
- they cannot protect the internal network from attacks which come via wireless networks with weak security.

## 9.8 Firewall categorization

Firewalls can be divided into following categories [81], [84]:

- (stateless) firewall with packet filtering (packet filter firewall),
- stateful firewall (stateful packet filter firewall),
- application gateway firewall (proxy server/firewall).

### **Packet filter firewall**

The packet filter firewall is the simplest type of firewalls. During operation they utilize information from all packets (that traverse them). This information relates to third and fourth layer of the OSI reference model. The main function of the packet filters is to compare some values from packet headers with predefined values in a configuration (that's why they are called static firewalls as well). These values (from packet headers) mainly cover:

- source and/or destination IP address,
- source and/or destination TCP / UDP port,
- transport protocol,
- TCP protocol flags: SYN, ACK, RST (only some implementations support them).

Advantages of the packet filters are their simplicity and speed because they read headers of only IP or TCP/UDP protocols. This operation is easy to execute and therefore the packet filter has lower requirements on a device performance. For users they appear to be transparent.

A disadvantage of the packet filters is that they are often vulnerable to attacks which e.g. spoof IP addresses, fragment packets to small packets (fragments), make use of a source routing (when a source specifies the packet path via the network in advance) or exploit vulnerability of application protocols or applications. The packet filters don't store information about a state of connectionless oriented communications (UDP and ICMP protocols). When a complex security policy is deployed configuration mistakes can appear.

Fig. 9.5 depicts several rules defined for packet filtering. This firewall has two interfaces (internal/inside and external/outside). The filtering rules are applied to the internal interface and based on them the firewall will realize e.g. following actions:

- a packet sent from the internal network to a DNS server in the external network is permitted by the firewall because the DNS protocol sends messages using the UDP protocol on port 53 (first rule).
- IP addresses from the internal network can ONLY contact a (external) web server with the IP address 147.232.22.82. The asterisk sign in a rule represents any source address (or port). This rule enables members of the internal network to spoof their IP address and to realize various attacks to the external networks.
- the third rule allows to any address from the internal network to contact any address from Internet on the destination TCP port 443 that belongs to HTTPS communications by default.
- any communication via DHCP protocol (that uses UDP and ports 67, 68) is permitted by the firewall (fourth rule).



- all other communications are forbidden (last rule) because all firewall rules are evaluated step by step starting by first row (upper row) then continuing with second row, and so on until a matching rule is found.

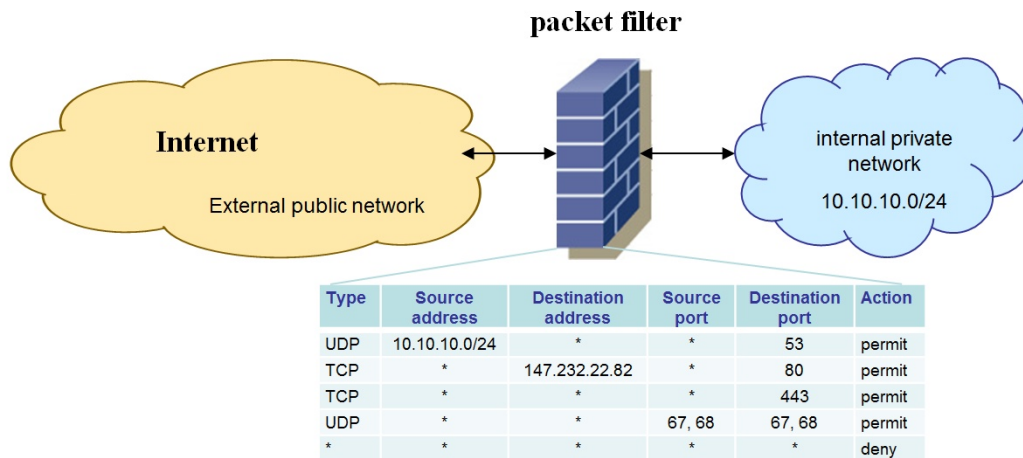


Figure 9.5 - The packet filter principle

### Stateful firewall (stateful packet filter)

In general, the stateful firewall works on the same layers of the ISO OSI model compared to the static (stateless) packet filters but they don't handle each packet independently. Besides the traffic filtering they monitor all communications which traverse them. It means that they also track packets' TCP/UDP ports (at TCP traffic, moreover, e.g. sequence numbers). Information about communication status is stored in a special dynamic state table. In other words the stateful firewall tries to identify a connection (flow) set up phase in outgoing traffic (from the internal network to the external network) that is permitted and information about it is stored in that state table. This information most often contains the source and destination IP address, source and destination port of an application protocol, protocol type, sequence numbers, etc.

Based on this information the firewall can identify those packets in incoming traffic which belong to permitted traffic and they are allowed to enter the internal network. Incoming packets which don't belong to any connection (initialized from the internal network) are dropped until a special rule exists that allows the firewall to let them in.

In case of TCP connections the firewall can identify a connection start and end by SYN and FIN flags in the TCP packet header. In case of connectionless protocols (UDP, ICMP) the firewall identifies a flow start by arrival of first flow packet and a flow end is set by timeout.

Disadvantages of the stateful firewalls are their higher requirements on hardware resources (higher processor power and memory space, etc.). Since they have to read and process more data from packet headers they introduce higher delays to the network than the packet filters. Compared to the packet filters they are also not able to identify an application layer protocol it means that if somebody sends a torrent file on TCP port 80 the stateful firewall will not block this traffic even though it is not allowed by the security policy.

### Application gateway firewall (proxy server/firewall)

As was already mentioned the packet filters (stateful as well as stateless) are not very efficient in application protocol filtering. For this purpose application gateway firewalls or so called proxy servers are applied which operate on behalf of their clients. They connect to the Internet on behalf their clients and transmit clients' data (and then they receive replies from the Internet and forward them to clients).

The application gateways work on all layers of the network model, i.e. they understand application protocols and can for example:

- to forbid flash videos from famous portals (e.g. youtube) but other content from those sites stays permitted.
- to monitor email communications and if the security policy doesn't permit attachments with ".doc" or ".pdf" files the application gateway can check email content and drop such attachments (even when they are compressed).
- to prohibit packet spoofing or illegal application usage on permitted ports (e.g. aforementioned torrent file via the port 80 would be dropped by the application gateway).
- to read encrypted communications.

On the other hand such deep inspection requires higher demands on device hardware power. As the application gateways temporarily store communication for processing they introduce higher delays to the network. Operations performed at the application layer can be realized only by software what brings additional load on a processor and memory. Some types of the application gateways require intervention (modification) on a client side as well.

### 9.8.1 Supportive educational materials

For this chapter following supportive materials and components are available:

- 3D VR application "Firewall",
- Manual for 3D VR application "Firewall".

## 9.9 Intrusion detection and prevention systems

The secured network access model based on deployment of the security application gateways is not able to secure the internal network against all types of attacks coming from the external network as well as against almost any attacks originating in the internal network. In order to increase a security level of the internal network it is necessary to deploy other system that tries to eliminate these shortcomings. For this purpose *intrusion detection systems (IDS)* and *intrusion prevention systems (IPS)* are implemented. These systems can work at a network level or directly in end devices (computers) as well. In general IDS collects and analyses information from various sources (within computers or a network) with aim to identify possible security violations including exploitations (attacks from within an organization) or intrusions (attacks coming from outside an organization) [81]. The IDS can eventually identify system vulnerabilities as well. For all of that, the IDS monitors all the incoming and outgoing network traffic or user behaviour on the end devices and it identifies suspicious conditions. Basic IDS functions are:

- monitoring and analysis of user and system activities,

- analysis of system configurations and vulnerabilities,
- evaluation of a system and file integrity,
- ability to recognize typical attack patterns,
- analysis of abnormal activity,
- tracking the security policy violations.

The IDS systems attempt to discover malicious activities (attacks) whereas the IPS systems attempt to protect the network and devices by dropping packets, forbidding access for packets, blocking connections, sending alarms to administrators, etc.

## 10. Content adaptation

When adapting the multimedia content, we can look at different aspects. Adaptation can be motivated by varying characteristics of the transmission path and specific display equipment, as well as the preferences or requirements of consumer.

The aim of the adaptation is generally to maximise the end to end quality of the content with regard to the possibilities of adaptation. From the point of view of the transmission path operator, can be required the maximum content quality at specified load of the network or its parts. One can also adapt to varying conditions of the transfer path: changing the error rate, changing latency, and so on. Or even change of the transmission path (and its provider) itself in the background - switch among a mobile connection, Wi-Fi, and a fixed connection. From the content manager perspective, he may want to minimize the storage size. From the user's perspective, he typically wants to maximise the quality of the content determined by him and adapted to the personal requirements and the equipment it uses. These requirements are often partly contradictory, have very many degrees of freedom to deal with and solutions are not straightforward.

### 10.1 Video quality and adaptation levels

The mere assessment of the quality of multimedia content is ambiguous for different users - is worse the video where is a better picture quality and worse audio quality? Or vice versa? And considering e.g. only audio part, what option is better, to slightly reduce the quality for a long time or massively reduce the quality for a shorter time? There are elaborated primarily the methods that evaluate the quality of each component of audio, static image, video separately. The basic division of these methods is depending on whether they work with the reference signal (original) or without it, whether they work objectively (mathematically established metrics) or subjective (using human observers). In the video e.g., objective methods with reference include VQM (ITU-T J. 147), PEVQ (ITU-T J. 246), among the subjective methods of the recommendations of ITU-R BT. 500 and ITU-T P. 910. These tests use the *Mean Opinion score (MOS)* metric, where observers evaluate the quality of the video under controlled conditions on a scale of 1 (poor) to 5 (good) and the results are averaged for the whole group of observers. The MOS metric was originally introduced to evaluate the quality of the talk signal in telecommunications (recommendation ITU-T P. 800.1). Some standardised objective methods have been created using mathematical models of the human observer behaviour (and the behaviour of the corresponding perception system), e.g. for audio, this is the **PEAQ** (*Perceptual Evaluation of Audio Quality*) method - Recommendation ITU-R BS. 1119. For video it is e.g. the above mentioned **PEVQ** (*Perceptual Evaluation of Video Quality*).

Since the requirements of an individual user may differ significantly from the average and we have possibility to adapt the content, the resulting quality could be significantly improved. E.g. when we have to reduce the channel transfer rate, some users may prefer

- to decrease the frame rate, preserving image sharpness and audio quality
- to preserve the frame rate, decrease the image quality, preserve the audio quality
- to turn off the audio and investing saved bits into the image quality

This, of course, may still be dependent on the device on which the content is rendered and even from the content itself. The user will likely assess the quality according to the different criteria for

- live transmission of the sport match
- action movie
- weather forecast in TV
- concert

When adapting the content, we know following three basic levels of adaptation:

- content selection - only relevant parts of multimedia content are selected (reflecting actual conditions, preferences, etc.)
- modality conversion - one component of a multimedia signal is converted to another component (e.g. text to Speech)
- content scaling/transcoding - is done e.g. by changing the format of the content to the format supported by the device and the content is delivered only in the needed resolution, and so on...

## 10.2 Content adaptation in static image and audio transmission

The need for content adaptation has resulted into standardisation of the file transfer formats that support the ability to have information stored in layers that enhance the basic layer (layer with basic information). One representative example is the JPEG standard for image compression. It specifies 2 modes that are suitable for usage in the content adaptation process (see Fig. 10.1):

- Progressive mode ([86], Annex G): The image information is encoded sequentially in that way, that the particular parts enhance quality in entire image. In decoder, we first have poor detail, but after receiving additional batches of information, the image is gradually improved. Moreover, the most important information, i.e. information that most diminish the reconstruction error is transmitted at the beginning and progresses to less important information. This can happen in two ways
  - spectral coefficient selection - spectral coefficients are transmitted in the full precision in the order that corresponds to their importance, from more important ones to the less important
  - successive approximation - spectral coefficients are transmitted using bit planes. Complete bit planes, one by one are transmitted in the order that corresponds to their importance, from more important to the less important ones.
- The hierarchical mode ([86], Annex J) allows better adaptation to the display device resolution. The image is transmitted in such a way that the information needed to reconstruct a smaller version of the image is transferred first, then is gradually transferred the information that will allow the image to be displayed in a high and larger resolution.

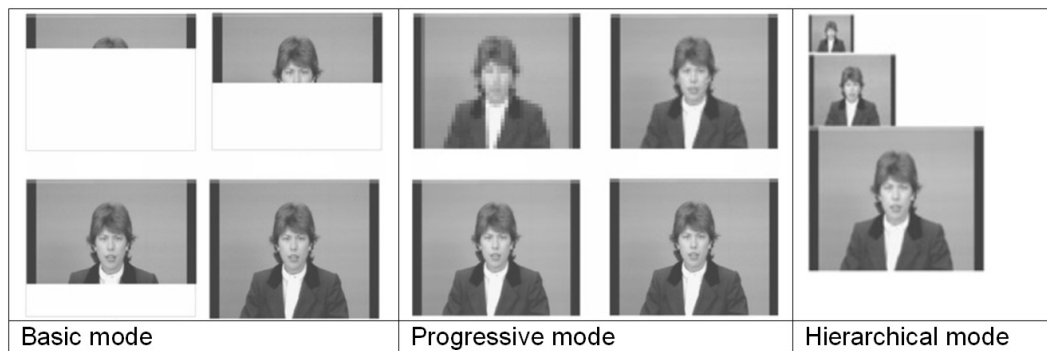


Figure 10.1 - Examples of gradually available information at the receiver during the transfer of a JPEG file when using different modes (the time axis for the displayed instants goes from left to right and then from top to bottom) [87]

The JPEG 2000 standard [88] for static image compression and transmission has introduced packetization for the information transmission, and in particular has introduced a concept of quality layers, which allows to directly structure data based on the contribution to quality. This subsequently allows stronger protection for the more important information and in the urgent case to throw away the less important information etc.

Currently, to the most effective standards for image compression belongs the **HEIF** (*High Efficiency Image File*) standard. It is actually part 12 of the MPEG-H recommendation (see below). This part uses algorithms for intra prediction from HEVC (see below), i.e. part of HEVC suitable for static image encoding). This standard is generally assumed to be an effective successor to the classic JPEG algorithm.

HEIF unlike JPEG 2000 (which was not well accepted by the consumer market) has received broad support, e.g. it is native part of the High Sierra OS. HEIF may contain several image items that represent the same picture, however, with a different spatial resolution, bit depth, gamut, encoding format, and profile.

In audio, the content adaptation was established especially in telecommunication systems. For example, **AMR** (*Adaptive Multi Rate*) codec (recommendation 3GPP TS 26.071) or its improved version **AMR WB** (*AMR Wide Band*) (recommendation G. 722.2 and 3GPP TS 26.190) encode the 20 ms portions of the speech signal and, depending on the current air interface conditions, version with appropriately strong protection is chosen. For good transmission conditions, that version will be chosen that contains maximal amount of useful information but with poor protection against errors. For bad transmission conditions, a version is chosen, that transfers little useful information, but is much more error resilient (see Fig. 10.2).

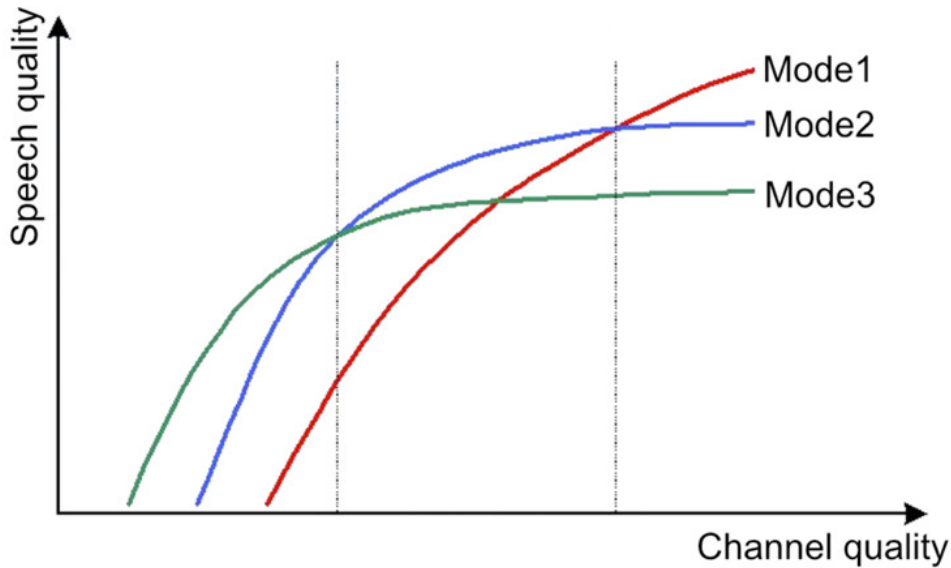


Figure 10.2 - AMR principle - For a given channel quality, a mode that provides the greatest quality of speech is selected, i.e. the best channel quality will be used Mode3 (the right area), when the quality is decreased, the mode Mode2 (middle area) shall be used and for lowest channel quality the Mode1 is appropriate.

### 10.3 Content adaptation in video transmission

In the video area, the notion of scalability was introduced by the MPEG-2 standard [89]. Here, the image stream is encoded in two layers, in first (basic) layer is transmitted the information, which can be further improved using enhancement layer, which contains additional information to achieve the full quality of the video. Some of the basic methods of scalability in the MPEG-2 are

- **SNR (Signal to Noise Ratio) based scaling:** an enhancement layer sends information that improves video quality at a given resolution
- **spatial scaling:** the base layer encodes video at lower resolution and enhancement layer encodes the additional information needed for full resolution

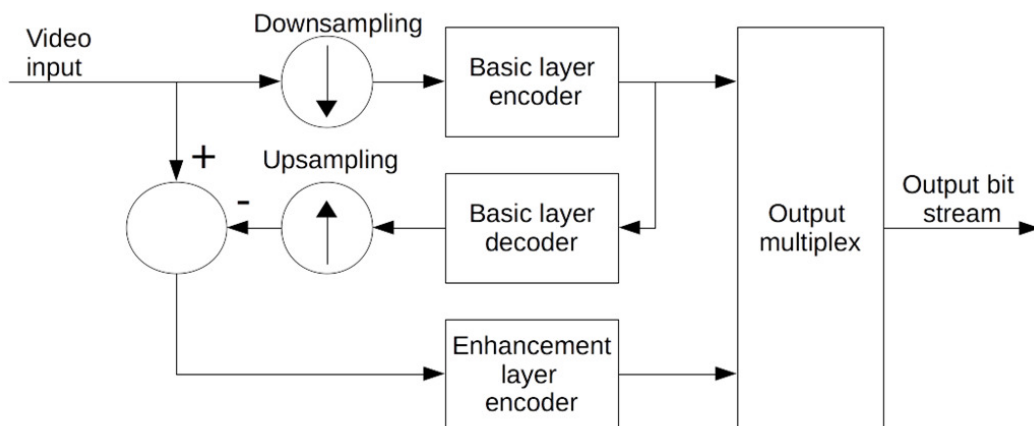


Figure 10.3 - Block diagram for MPEG-2 spatial scaling encoding.

For spatial scaling (see Fig. 10.3), the input video signal is spatially-reduced horizontally and vertically. Reduced signal is then standardly encoded by the basic layer encoder. Then it immediately decoded and enlarged to the original dimension. The differential image is encoded using the encoder's enhancement layer. Finally, the outputs from the basic and enhancement layers are multiplexed to the output bit stream.

The MPEG-2 standard does not support audio scalability. This introduces the MPEG-4 standard in the form of MPEG-4 SLS (Scalable lossless Coding), which allows you to encode residual information as an additional layer and achieve the quality of lossless coding. MPEG-4 also extends the scalability of MPEG-2 image information by allowing to use more than one enhancement layer. In addition to the spatial and SNR scaling, there is also introduced the time scaling within the MPEG-4 AVC/H. 264 [90] standard. Time scaling means, that the pictures are grouped in time into multiple sequences in such way, that when needed some sequence can be dropped, resulting in a reduction in the frame rate. In appendix G of the MPEG-4 AVC standard under the name JRC (Scalable Video Coding) are given additional recommendations how to use the scalability effectively [91].

The H.265 coder - standardized in MPEG-H Part 2 as **HEVC** (*High Efficiency video coding*) [92] delivers scalability for videos up to a resolution of 8K, i.e. 8192x4320 points and up tenths of percent improves the compression rate compared to MPEG-4 AVC. Time scalability was already present in the first version of the standard, but spatial, SNR scalability and scalability based on bit depth and gamut and their combinations were added only in the second version of the standard in the form of **SHVC** (*Scalable HEVC extensions*) in October 2014. At the same time, there were added options for efficient transfer of 3D video (stereoscopic, with a depth map) and *MultiView (MV)* video.

The MPEG21 standard, which was finalised in 2003 and defines an open framework for multimedia applications, is in part 7 dedicated to the issue of adapting content in the form of *Digital Item Adaptation (DIA)* [93]. The term Digital item includes some multimedia content, along with all the necessary metadata. DIA specifies a set of descriptive tools to optimize the adaptation of multimedia content. Specifies how the environment should be described in the adaptation, describing the characteristics of the network, equipment, user preferences and user's environment:

- device capabilities: available codecs, I/O options (display, audio, microphone, ...)
- network characteristics: Network capabilities (max. capacity, minimum guaranteed bandwidth, ...), delay, jitter, error rate, available bandwidth
- user characteristics: user preferences, usage history, general user information, preferences for presenting individual modalities - e.g. preferred audio volume, equaliser setting, image resolution, colour temperature, saturation, contrast.
- accessibility: user characteristics for users with disabilities e.g. visual system (e.g. degree and type of colour vision disorder), auditory system (hearing limitation at different frequencies)
- physical environment of the user: location and time of use, audio-visual properties of the environment (e.g. level and characteristics of noise), lighting conditions

The standard also covers the scalable content in the part MPEG-4 SVC and provides a generic tool to describe the syntax of a bit stream and its use for creation of an adapted version of the



content. In addition to using scalability, it also describes the options for converting content modalities (modality translation), e.g. audio to text and vice versa.

In 2012, there was standardized the MPEG-DASH (*Dynamic Adaptive Streaming over HTTP*) - ISO/IEC 23009-1, which describes the client's interaction with server in the adaptive delivery of content via the HTTP protocol [94]. MPEG-DASH is agnostic from codec, it can be used with any modern encoding procedure for encoding video, such as H.264, H.265, VP9, etc. The principle is, that on the server side the content is divided into segments, while each segment is encoded using various bit rates using variety of parameter settings. Information about segments and their versions is available through the **MPD** (*Media Presentation Descriptor*) file. The client device will first download the MPD file, then starts downloading the video by downloading that segments versions that best correspond to the current conditions on the transfer channel (see Fig. 10.4). The obvious criterion is to manage to download in time the highest quality or the most appropriate version for the consumer.

MPEG-DASH is used for example in Netflix and Youtube when distributing the content, even when streaming live video. There are systems that can effectively use the MPEG-4 SVC scalability when using MPEG-DASH (e.g. [95]). Data samples (MPD files, video files, etc.) using the SVC are provided for example in [96].

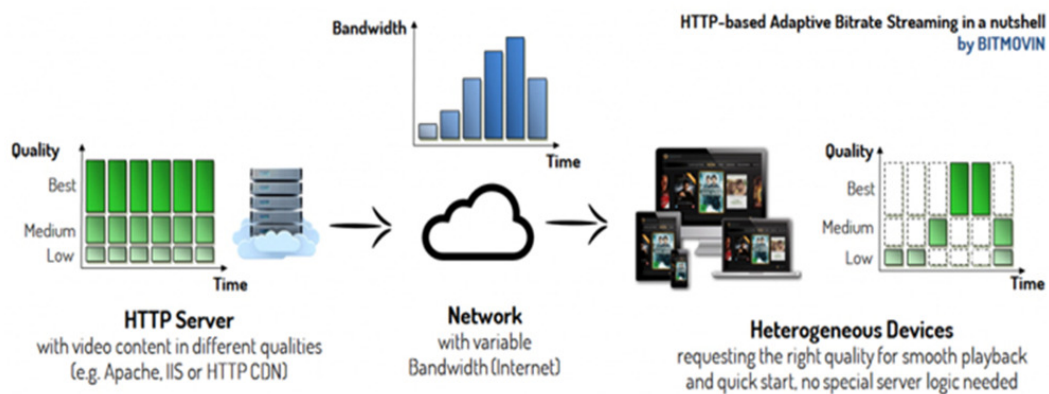


Figure 10.4 - The principle of adaptive multimedia delivery using MPEG-DASH [97]

In the context of MPEG-DASH the question arises, why the client, when there are good transmission conditions does not download the entire content as soon as possible? (Otherwise maybe the conditions get worse and client will be maybe forced to fetch the worse quality to get the content in time.) The answers are several, e.g.:

- limitation of available client resources (memory),
- why to download the data that the user will maybe not have interest to see,
- why to generate network load at a time when they may be more urgent needs of others.

MPEG-DASH thus allows the content delivery control under different transport channel conditions and to optimize the quality and transmitted volume of data and other criteria (e.g. the battery life of the device) from the point of view of individual users. Into optimization of the decision-making in the client (which MPEG-DASH standard does not standardise) it is currently put a lot of effort, for example [98].

In 5G communication networks there is an emphasis not only on the increasing capacity of the network, but also on the growing intelligence of the network to further increase this capacity [99]. In this context, several network extensions have been proposed, that comparing to MPEG-DASH introduce the intelligence on the network side, e.g. using the cache (cache) placed as close as possible to users [100] [101]. ETSI calls this component the **MEC** (*Mobile Edge Computing*) [102] and typically locates it on the edge of the mobile network, i.e. for example in NODE-B. Using such component, the load in the rest of the network can be reduced, as only part of the access network between the cache and the end-users is more loaded [103].

For MPEG-DASH there exist currently the research and standardisation efforts to extend the adaptation for the delivery of VR content (DASH-VR). The main motivation is based on the fact that usually, when watching 360° videos, we see at once only about 1/12 of the total image information. The rest is beyond our **FOV** (*Field Of View*). There are 3 levels foreseen [104]:

- Basic level - delivering the entire 360° video
- **SRD** (*spatial relationship description*) - Provides a partitioning technique (tiles) so that the user only receives the content that is he currently viewing plus some area around
- SRD combined with SVHC - when combining SRD with SHVC, one can efficiently scale the quality for each area when splitting a video into spatial areas. An example of such scaling is given in Fig. 10.5.

For the implementation of the adaptation on the network-side, it is possible to use the architecture of **SDN/NFV** (*Software-Defined networking/Network Function virtualization*) networks. In its modules **VNF** (*Virtualised network functions*), which may be in the position of the **EC** (*Edge computing*), i.e. they may provide the computing power at the network boundary, it is possible to efficiently build adaptation on the network side in 5-generation networks.

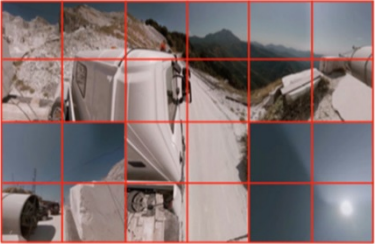


Example of splitting a video to areas and their versions on the server side		Selection of areas delivered to client, based on actual FOV of the user
Areas in full resolution	Areas in half resolution	
		

Figure 10.5 - Example for the VR content delivery, SRD combined with SVHC, available areas and their versions and selection of areas delivered to client, based on actual FOV of the user [105]

# 11. 5G Network architectures, services and applications

## 11.1 Introduction

5G is an abbreviation for the 5th generation of wireless system, which represents the next evolutionary stage in the development of mobile networks. Requirements for 5G networks were specified by the *International Telecommunication Union (ITU)* under IMT-2020 [106] and are being developed within **3GPP** (The *3rd Generation Partnership Project*) in Release 15 (5G Phase 1) [107] and Release 16 (5G Phase 2) [108].

The 5G system will be able to provide the widest range of services and applications in the history of mobile and wireless communications categorized under *enhanced mobile broadband (eMBB)*, *massive machine-type communications (mMTC)* and *ultra-reliable low-latency communications (URLLC)*.

The ITU defines the following requirements for 5G networks [106], [109]:

- 1000x greater volume of mobile data to cover a certain area (e.g. 0.75 Tbit/s for stadiums)
- 1000x greater number of connected devices (max. density  $\geq 1$  million terminals/km<sup>2</sup>)
- 100 times higher transmission speeds (peak  $\geq 1$  Gbit/s for office applications)
- 1/5x end-to-end delay (target:  $\leq 1$  ms)
- Transfer rate guaranteed to the user  $\geq 50$  Mbit/s
- Support for IoT terminals  $\geq 1$  trillion ( $10^{12}$ )
- Availability  $\geq 99.999\%$  (for some specific services)
- Promoting mobility at a speed of  $\geq 500$  km/h for land transport

Performance requirements for high data rate and traffic density scenarios as well as for low-latency and high-reliability scenarios can be found in [110].

To meet these requirements, the 5G network also brings improvements in network architecture. In addition to supporting multiple access technologies, it also provides modular and flexible network architecture. 5G network will involve the integration of several cross-domain networks, and the 5G systems will be built to enable logical network slices across multiple domains and technologies to create tenant- or service-specific networks. The network slicing will be designed from an end-to-end perspective spanning over different technical domains (e.g. core, transport and access networks) and administrative domains (e.g. different mobile network operators) including management and orchestration plane.

The security architecture will be natively integrated into the overall architecture, e.g. to ensure the requirements of the enhanced applications and services pertaining to the safety-critical use cases.

## 11.2 5G system architecture

5G system architecture is defined to support data connectivity and services enabling deployments to use techniques such as e.g. **NFV** (*Network Function Virtualization*) and **SDN** (*Software Defined Networking*).

The main features of 5G system architecture are:

- Separates the *User Plane (UP)* functions from the *Control Plane (CP)* functions that allow scalability, evolution and flexible deployments.
- Modularize the function design, e.g. to enable flexible and efficient network slicing and define procedures (i.e. the set of interactions between network functions) as services, so that their re-use is possible.
- Enable each network function to interact with other network function directly if required. The architecture does not preclude the use of an intermediate function to help route control plane messages.
- Minimize dependencies between the access network and the core network. The architecture is defined with a converged core network with a common access network - core network interface which integrates different access types.
- Support concurrent access to local and centralized services. To support low latency services and access to local data networks, user plane functions can be deployed close to the access network.

The 5G architecture is defined as service-based and the interaction between network functions is represented in two ways.

- **Reference point representation** (Figure 11.1) - shows the interactions that exist between the network function services in the network functions described by point-to-point reference point (e.g. N4) between any two network functions (e.g. UPF and SMF).
- **Service-based representation** (Figure 11.2) - where network functions (e.g. AMF) within the control plane enables other authorized network functions to access their services.

Network functions within the 5G control plane shall only use service-based interfaces for their interactions.

The 5G System architecture consists of the following network functions:

- **UE** (*User Equipment*), **R(AN)** (*(Radio) Access Network*), **UPF** (*User Plane Function*), **DN** (*Data Network*),
- **AMF** (*Access and Mobility Management Function*), **SMF** (*Session Management Function*),
- **NSSF** (*Network Slice Selection Function*), **NEF** (*Network Exposure Function*), **NRF** (*Network Repository Function*), **AUSF** (*Authentication Server Function*), **UDM** (*Unified Data Management*), **PCF** (*Policy Control Function*), **AF** (*Application Function*), **UDSF**

(Unstructured Data Storage Function), **UDR** (Unified Data Repository), **5G-EIR** (5G-Equipment Identity Register).

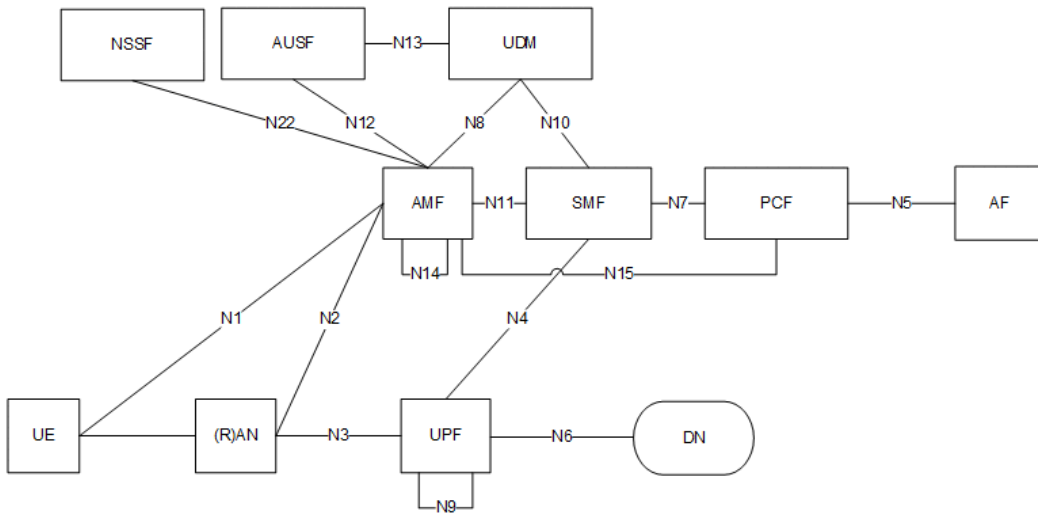


Figure 11.1 - 5G system architecture - reference point representation [111]

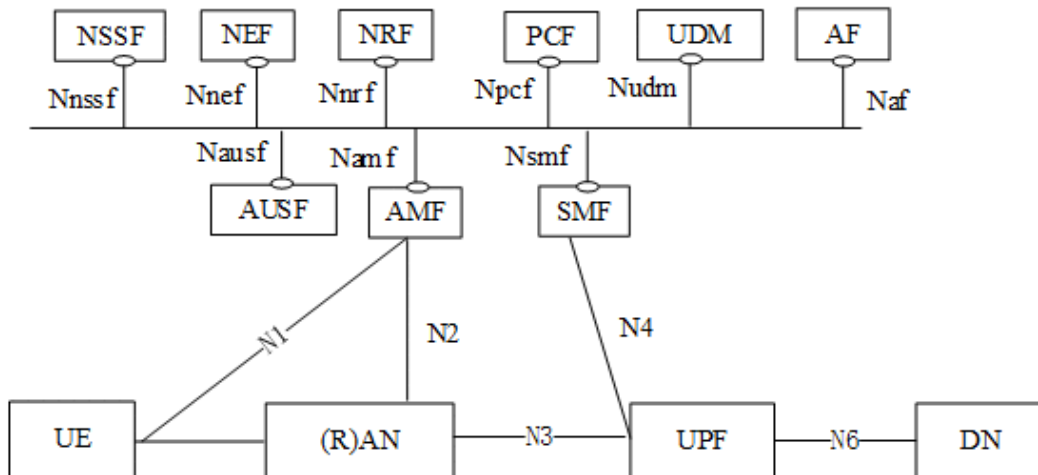


Figure 11.2 - 5G system architecture - the reference point representation [111]

A network function can be implemented either as a network element on a dedicated hardware, as a software instance running on a dedicated hardware, or as a virtualized function instantiated on an appropriate platform, e.g. on a cloud infrastructure

The more detail description of network functions can be found in [111], [112], and [113].

### 11.2.1 Network slicing

One of the major novelties of the 5G network is a network slicing aims for building dedicated logical networks that exhibit functional architectures customized to the respective telco services, e.g., eMBB, URLLC, mMTC, **V2X** (*Vehicle-to-everything*) etc.

Network slice is a logical network serving a defined business purpose or customer, consisting of all required network resources configured together. It is created, changed and removed by management functions. To support network slicing, the management plane creates a group of network resources. It connects with the physical and virtual network and service functions as appropriate, and it instantiates all of the network and service functions assigned to the slice. For slice operations, the control plane control all the network resources, network functions, and service functions assigned to the slice. It configures them as appropriate, in order to provide an end-to-end service. In particular, ingress routers are configured so that the appropriate traffic is bound to the relevant slice.

The creation of slices can be:

- **business-driven** - slices are created to support different types and service characteristics or business cases,
- **technology-driven** - slices are created by grouping of physical or virtual resources (network, compute, storage) which can act as a sub network or a cloud.

From a business point of view, the slice includes a combination of all the relevant network resources, network functions and service functions (physical or virtual) in all of the network segments (access, core and edge) required to fulfil a specific business case or service, including **OSS** (*Operations Support System*) and **BSS** (*Business Support System*).

The principle network slicing is illustrated in Figure 11.3.

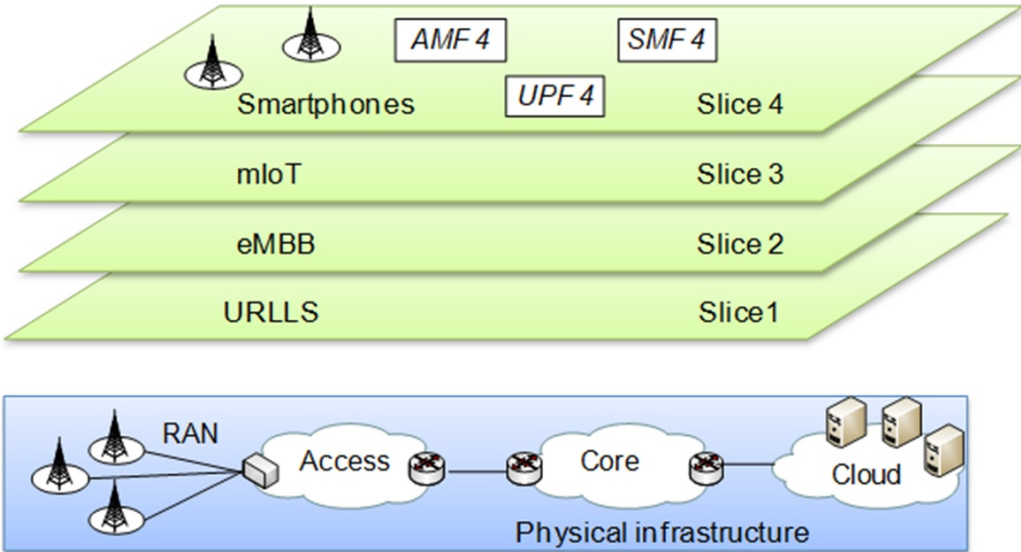


Figure 11-3 Network slicing

Note: Due to the complexity of network slicing and actual state of the technology, the first 5G networks will use 5G RANs in combination with existing **EPC** (*Evolved Packet Core*) networks.

## 11.2.2 RAN (Radio Access Network)

The *Radio Access Network (RAN)* is an important part of the 5G network which allows achieving high bandwidth in the 5G networks.

Like the rest of the network, the 5G RAN brings some improvements over previous mobile network versions. In addition to using different frequencies and code schemes, it is also the use of cloud technology.

The 5G RAN provides the so-called RAN real time functions a RAN non-real time functions.

The RAN *real time (RT)* functions include access network scheduling, link adaptation, power control, interference coordination, retransmission, modulation, and coding. These functions require high real-time performance and computing load. The deployment of sites must include dedicated hardware with high accelerator processing specifications and performance, whilst located in close proximity to services.

The RAN *non-real time (NRT)* functions include intercell handover, cell selection and reselection, user-plane encryption, and multiple connection convergence. These functions require minimal real-time performance, latency requirements to tens of milliseconds and are suitable for centralized deployment. Universal processors can be deployed in a **C-BBU** (*Cloud Baseband Unit*) pool or site according to vast service requirements.

Cloud RAN architecture is used on the RAN side to implement RAN real time functions, on-demand deployment of nonreal time resources, component-based functions, flexible coordination and RAN slicing.

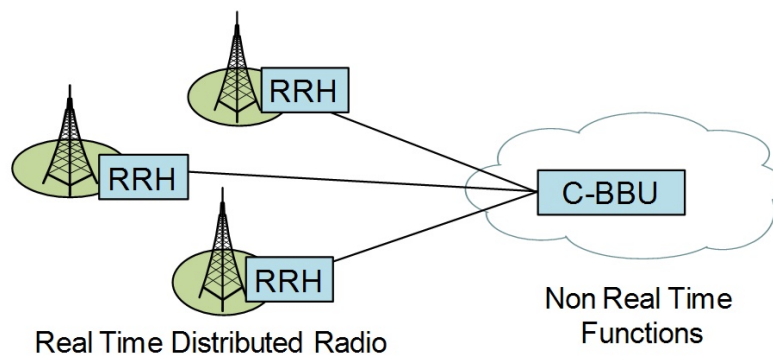


Figure 11.4 - Cloud RAN architecture

## 11.2.3 Core Network

The transport network consists of SDN controllers and underlying forwarding nodes. SDN controllers generate a series of specific data forwarding paths based on network topology and service requirements. The enabling plane abstracts and analyses network capabilities to implement network optimization or open network capabilities in the form of API.

Networks implement policy control using dynamic policy, semi-static user, and static network data stored in the unified database on the core network side.

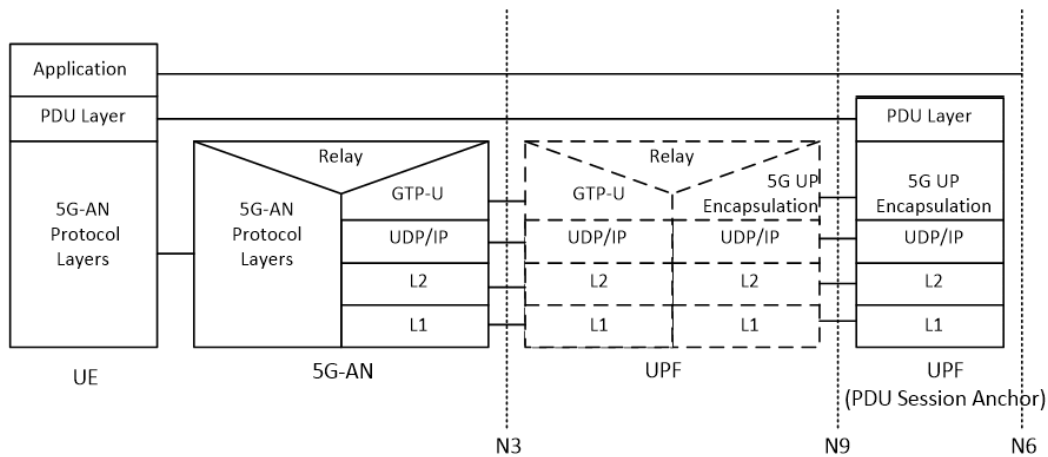


Figure 11.5 - User plane protocol stack [111]

Figure 11.5 illustrates the protocol stack for the User Plane transport related with a **PDU (Protocol Data Unit) Session**. The **PDU layer** corresponds to the PDU carried between the UE and the DN over the PDU Session. When the PDU Session Type is IPV6, it corresponds to IPv6 packets; When the PDU Session Type is Ethernet, it corresponds to Ethernet frames; etc. **GTP-U (GPRS Tunnelling Protocol for User plane)** [114] supports multiplexing traffic of different PDU Sessions by tunnelling user data over N3 interface between the 5G-AN node and the UPF in the backbone network. GTP will encapsulate all end user PDUs. **5G UP Encapsulation:** This layer supports multiplexing traffic of different PDU Sessions over N9 (i.e. between different UPF of the 5G core network). It provides encapsulation on a per PDU Session level.

#### 11.2.4 Management and Orchestration

The top layer of the network architecture implements end-to-end automatic slicing and network resource management. Component-based control planes and programmable user planes allow for network function orchestration to ensure that networks can select corresponding control-plane or user-plane functions according to different service requirements.

The management and orchestration of 5G networks is based on the advanced ETSI / NFV **MANO (Management and Orchestration)** architecture [115]. Special aspect can be considered as network slicing, multi-tenancy, multi-domain or multi-operator operation.



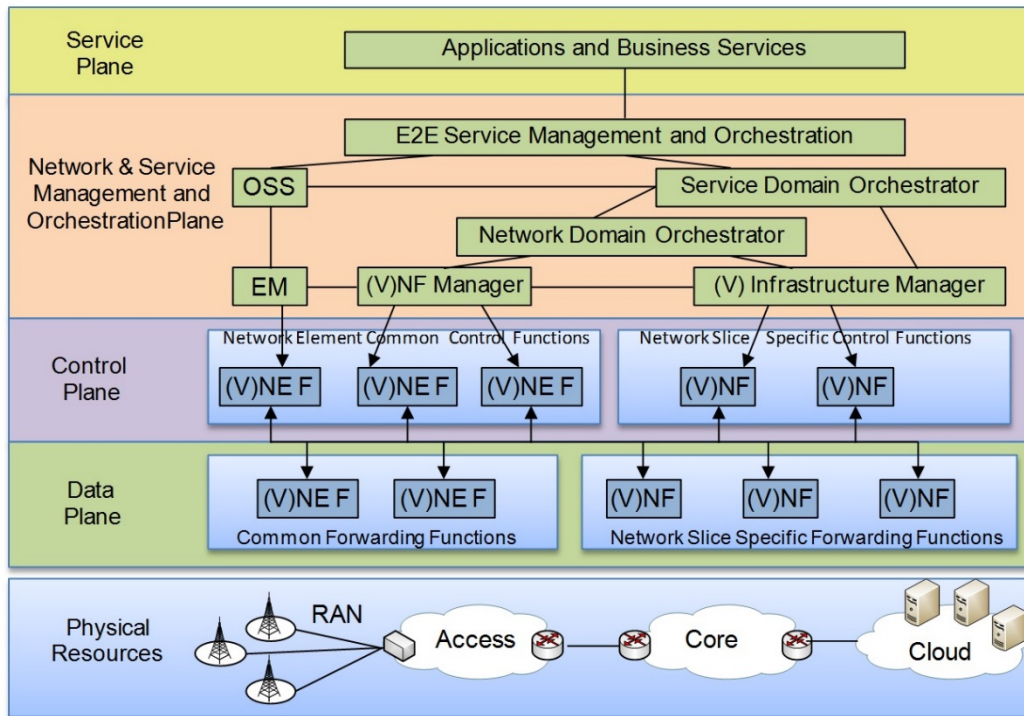


Figure 11.6 - Management and orchestration of the 5G network

### 11.2.5 Security

The 5G security concept is based on use of domains. A **Domain** is a grouping of network entities according to physical or logical aspects that are relevant for the 5G network [116]. There are three different types of domains:

- **Infrastructure Domains** - focus on the relevant physical network aspects, i.e. they contain the “hardware” in the network.
- **Tenant Domains** - are logical domains executing in infrastructure domains.
- **Compound Domains** - consist of a collection of other domains, grouped together according to some 5G relevant aspects, e.g. ownership. One important type of compound domains are **Slice Domains**.

Another important element of the security architecture is a **Stratum**. The Stratum is a grouping of protocols, data, and functions related to one aspect of the services provided by one or several domains.

- **Application Stratum** represents the application process provided to the end-user.
- **Home Stratum** contains the protocols and functions related to the handling and storage of subscription data and home network specific services.
- **Serving Stratum** consists of protocols and functions to route and forward data/information, user or network generated, from source to destination.
- **Transport Stratum** supports the transport of user data and network control signalling from other strata through the network.

- **Access Stratum** is a sub-stratum of the Transport Stratum. It is located between the edge node of the serving network domain and the UE Domain.
- **Management Stratum** comprises aspects related to conventional network management (configuration, software upgrades, user account management, log collection/analysis, etc.) and, in particular, security management aspects (security monitoring audit, key and certificate management, etc.).

The last element of the security architecture is *Security control classes* [117]. The Security control classes include Authentication, Confidentiality, Integrity etc.

### 11.3 5G services

Large number of new use cases to be supported by 5G networks have been described and analysed by 3GPP and ITU-T. Many supported use cases are variations of a small set of basic service classes offered by 5G networks will be:

- *Enhanced mobile broadband (eMBB)* - that will support gigabits of bandwidth for 4K video, stereo 360° video, *virtual reality (VR)*, etc.
- *Massive machine-type communication (mMTC)* - that will be able to connect billions of sensors and machines together with densities of up to 200 thousand per square kilometre, making it possible to support smart cities.
- *Ultra-reliable, low latency communication (uRLLC)* - that will allow for instant feedback with high reliability and will enable e.g. real-time remote surgeries, remote control over robots in manufacturing and autonomous driving [117].

#### Legacy services

5G networks will also offer legacy services, i.e. services provided by previous generations of mobile networks. Some of them are:

- **IMS (IP Multimedia Core Network Subsystem) support** - The IP multimedia subsystem includes all core network elements for delivering multimedia services. IP multimedia services are using **IETF (Internet Engineering Task Force)** protocols wherever possible utilising IP Connectivity Access Networks. The IP multimedia core network subsystem enables operators to offer multimedia services, e.g. voice, video, messaging, data and web-based technologies [118].
- **SMS - Short message service** will be realized as SMS over **NAS (Network-attached storage)** [112].
- **Public Warning System** - The cell broadcast service can be used to transfer **CBS (cell broadcast service)** messages related to public warning. Warning message delivery is similar to cell broadcast service [119].
- **Location services** - Location services identify and report current location information of user using multiple positioning methods (e.g. GPS, GLONASS...). This location information consists of geographic location (geographical coordinates) and velocity (speed and direction of user equipment) [120].

## 11.4 5G applications

Previous generations of mobile networks were increasing bit-rate and decreasing latency dealing with massive mobile broadband. Technology of 5G networks will focus on both high bit-rate and low latency, as well as low bit-rate devices and sensors for **IoT** (*Internet of things*) applications. Figure 11.7 summarizes main applications that will be supported in 5G networks with their required bitrates and latencies.

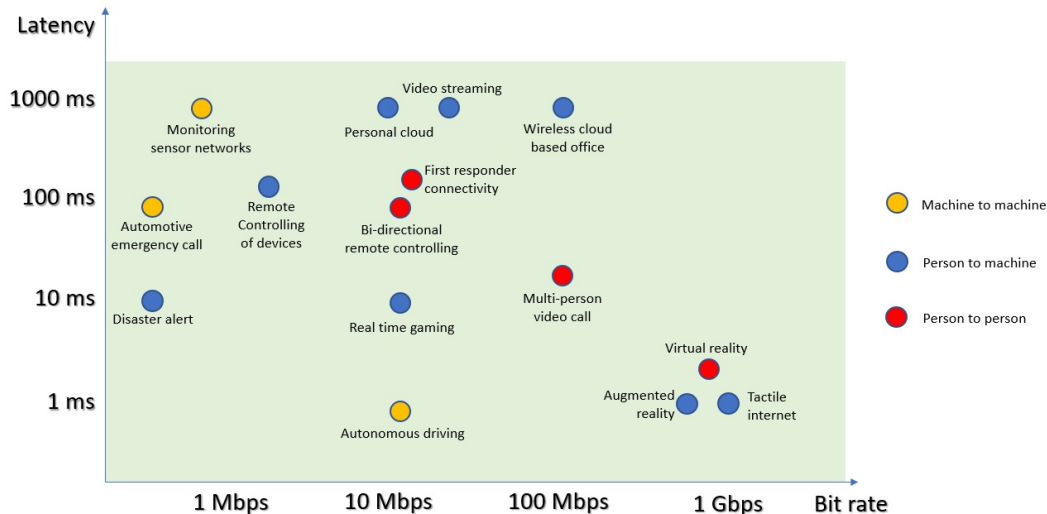


Figure 11.7 - Future applications to be supported by 5G networks [121]

### Broadband experience

Thanks to enhanced mobile broadband, 5G networks will provide better connectivity both for indoor and outdoor consumers with faster data transfers and with lower latency even in crowded areas, on events, public transportation and football stadiums.

### Media everywhere

Massive increase in mobile broadband demand in the last few years is caused by media consumption. 5G will provide the capability to stream Live TV, On-demand video streaming and gaming or shopping in AR/VR. The media will be provided in 4k, 8k, **HDR** (*high dynamic range*) and **HFR** (*high frame rate*) quality, 3D 360° video with haptic feedback.

- **Online VR Games** - virtual reality is a scenario that simulates a realistic experience generated by a computer with real time interaction with objects and other players. Online virtual reality will need reliable and low latency **MTP** (*motion to photon* latency) below 20 ms to convince the user's mind that he is present in another place to prevent dizziness. Other crucial parts of AR (augmented reality) and VR experience are the display resolution and video data bitrate. For low resolution 360° VR experience is at least 25 Mbit/s throughput is required. The required throughput for VR experience comparable to HDTV is up to 80-100 Mbit/s. Ultimate VR quality is 8K\*8K resolution with frame rate of 120 **FPS** (*frames per second*) requiring data transfer speed of up to 1 Gbit/s [122].
- **3D 360° video** - 360 video captures scene in every direction, while standard monoscopic video is displayed on a sphere, stereoscopic (3D) video adds another level of immersion

- depth between background and foreground. Typical formatting of 360° video is monoscopic or stereoscopic video stream using equirectangular projection stitched from more scene cameras. For comparison, 360° video needs typically 3-5 higher bitrate than normal, single view video, while 3D video typically requires 30-60 % more throughput than monoscopic video [123].

## **Internet of things (IoT)**

Internet of things is the network of physical devices (e.g. temperature monitoring, home appliances, wearables, vehicles and other items). These devices are embedded with electronics, sensors and software for control as well as connectivity which enables them to exchange data. 5G networks will support density of connecting devices up to 200 000 per square kilometre (massive machine type communication). IoT will allow objects to be remotely discovered and controlled through network infrastructure [124]. This encompasses technologies such as:

- **Smart vehicles** - In the recent years there is a lot of focus to provide autonomous driving experience. Developing connected cars and other smart vehicles is the focus of various companies such as Uber, Google and Volvo. 5G will provide the network part of this technology. It is expected that by the year of 2020, there will be 10 million self-driving cars on the road and more than 250 million smart vehicles (cars connected to 5G network) sharing the road with them. Smart vehicles will be working together with both navigation and broadcast satellites. As part of development of smart cities, smart vehicles will also communicate with smart phones and roadside units, making them important component of IoT. This interconnection will enable cars to alert drivers in case of a chain-reaction collision just a few cars ahead of them, or when someone is about to drive through a red light [125]. The new technologies will enable cars to send drivers alerts for such things as a chain-reaction collision a few cars up the road or when someone is about to drive through a red light. It will also send information about vehicles braking ahead and vehicles in blind spots real-time. The information about other cars and the environment around them will be sent anonymously as *vehicle-to-vehicle (V2V)* and *vehicle-to-infrastructure (V2I)* communication. Many modern cars have implemented instruments using radar or ultrasound detection of obstacles or vehicles, however the range of these sensors is limited only to a few tens of meters around the vehicle and cannot see past vehicles surroundings. Driver alerts through the network will be quick and intuitive, using lights and audio to warn other cars of potential dangers. Broadcasted information from other vehicles will be processed with on-board computers in vehicles and each time chance of potential collisions is calculated. This technology could prevent lot of accidents and fatalities, while also reducing the number of hours spend in traffic [126].
- **Smart Cities** - it is necessary that a city takes responsibility for its digital infrastructure similarly to previously taking responsibility for analogue infrastructure and its services, and ensured that these services worked efficiently together. This will provide means for the city to be in control of certain areas it is responsible for and their development. Smart cities will have a huge demand on data throughput in multiple sectors such as public transportation, emergency services and water management system. Integration of these sectors with IT system will lead to better planning and real-time operation, which will result in the automated responses to changing conditions. Smart City applications need to be integrated with sensors connected to networks. It is necessary to have uniform process for connectivity of these sensors and a network that is capable of supporting multiple data flows with multiple varying performance and functionality requirements [127].

- **Smart homes** - A smart home is a home setup using technology that assists and automates everyday activities and allows to remotely control smart devices that are parts of homes from anywhere. Generally, these devices can be accessed through one central device - a smartphone, a tablet or a laptop which is owned by almost everyone these days. With these devices we can control the security access to the home, settings on thermostat, or lower the shades. Smart home and its devices are interconnected through the Internet and they can exchange information. Smart home setup can be designed effectively to change its power consumption and save lot of time and money [128].
- **Smart postal** - Smart postal implemented as the IoT will help reduce visits to empty mailboxes. The new mailboxes will use sensors that will be connected to an **NB-IoT** (*narrow band internet of things*) network to provide real-time information on their status. Using this real-time information, the mailboxes will be emptied according to need, which could prove useful during peak seasons. Sensors will be sending data directly to reporting system, which will help monitor movements of mail centrally and in real time. This will allow for each mail carrier and driver to receive information relevant to his function [129].
- **Smart wearables** - Wearable devices, such as fitness bands or smart watches are becoming part of everyday life. These devices are incorporating electronics and software with sensors and wireless connectivity. There are an increasing number of use cases for these devices in various sectors, such as household, health and construction [121]. Nowadays these devices are used mainly for entertainment and monitoring physical activity, but their use is also expanding in healthcare applications (e.g. wearable **ECG** (*electrocardiograph*), *EEG* (*electroencephalograph*), blood pressure, pulse, respiration, sleep and motion monitoring). Smart wearables will be clinically useful for monitoring of real-time, long-term physiological and pathological processes. This could help to better understand chronic illnesses such as cardiovascular disease and sleep disorder. The research in smart wearables is based on clinical big data and artificial intelligence to build more robust techniques to identify a disease [130].
- **Critical control of remote devices** - One of the main goals and upgrades of 5G networks will be providing better connectivity for heavy machineries, smart-grids and remote surgeries. This will help mostly to manufacturers, miners and healthcare industry. Critical control includes mainly a machine-to-machine communication, for which high reliability and low latency are essential factors. This includes robot-to-robot communication in factories and robots in the field (e.g. harvesting apples [131]), as well as communication between drones [132]. *Ultra-reliable, low latency communication* (**uRLLC**) will be crucial for example for remote surgeries [133].

## LITERATURE

- [1] 3GPP TS 29.163; V6.7.0 „Interworking between the IP Multimedia (IM) Core Network (CN) subsystem and Circuit Switched (CS) networks“ (2005-06)
- [2] Mikóczy, E., Podhradský, P.: Evolution of IPTV Architecture and Services towards NGN, in Recent Advantages in Multimedia Signal processing and Communications, Springer-Verlag, 2009, pp. 315-339
- [3] Dúha, J., Galajda, P., Kotuliak, I., Levický, D., Marchevský, S., Mikóczy, E., Podhradský, P. at al.: Multimedia ICT technologies network platforms and multimedia services, Vydal: STU Bratislava, 2005, ISBN 80-227-2310-X
- [4] Ferkl, L., Šmejkal, L., Sládek, O., Podhradský, P., Dúha, J.: Teleinformatics in Industrial Automation, LdV ELeFANTC, Educational publication, Vydal: AGROGENOFOND Nitra, 2007, ISBN 978-80-89240-14-2
- [5] ITU-T Recommendation X.200 (11/93) [ISO/IEC 7498-1:1994], Open Systems Interconnection - Basic Reference Model.
- [6] Postel, J.: RFC 791 - Internet Protocol, IETF, September 1981.
- [7] Deering, S., Hinden, R.: RFC 2460 - Internet Protocol, Version 6 (IPv6) Specification, IETF, December 1998.
- [8] Postel, J.: RFC 792 - Internet Control Message Protocol, IETF, September 1981.
- [9] Conta, A., Deering, S., Gupta, M.: RFC 4443 - Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification, IETF, March 2006.
- [10] Deering, S., E.: RFC 1112 - Host extensions for IP multicasting, IETF, August 1989.
- [11] Postel, J.: RFC 793 - Transmission Control Protocol, IETF, September 1981.
- [12] Postel, J.: RFC 768 - User Datagram Protocol, IETF, August 1980.
- [13] Stewart, R., Ed.: RFC 4960 - Stream Control Transmission Protocol, IETF, September 2007.
- [14] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., Jamin. S.: RFC 2205 - Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification, IETF, September 1997.
- [15] Droms, R.: RFC2131 - Dynamic Host Configuration Protocol, IETF, March 1997.
- [16] Mockapetris, P.: RFC 1034 - Domain names - concepts and facilities, IETF, 1987.
- [17] Dierks, T., Rescorla, E.: RFC 5246 - The Transport Layer Security (TLS) Protocol Version 1.2, IETF, August 2008.

- [18] Freier, A., Karlton, P., Kocher, P.: RFC 6101 - The Secure Sockets Layer (SSL) Protocol Version 3.0, IETF, August 2011.
- [19] Postel, J., Reynolds. J.: RFC 854 - Telnet Protocol Specification, IETF, May 1983.
- [20] Postel, J., Reynolds. J.: RFC 959 - File Transfer Protocol, IETF, October 1985.
- [21] Ford-Hutchinson, P.: RFC4217 - Securing FTP with TLS, IETF, October 2005.
- [22] Fielding. R., T., Gettys, J., Mogul, J., C., Nielsen, H., F., Masinter, L., Leach, P., J., Berners-Lee, T.: RFC2616 - Hypertext Transfer Protocol -- HTTP/1.1, IETF, June 1999.
- [23] Berners-Lee, T., Fielding, R., T., Masinter, L.: RFC3986 - Uniform Resource Identifier (URI): Generic Syntax, IETF, January 2005.
- [24] Myers, J., G., Rose, M., T.: RFC 1939 - Post Office Protocol - Version 3, IETF, May 1996.
- [25] Crispin, M.: RFC 3501 - Internet Message Access Protocol - Version 4rev1, IETF, March 2003.
- [26] Klensin, J.: RFC 5321 - Simple Mail Transfer Protocol, IETF, October 2008.
- [27] Saint-Andre, P.: RFC 6120 - Extensible Messaging and Presence Protocol (XMPP): Core, IETF, March 2011.
- [28] Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation, November 2008.
- [29] SOAP Version 1.2 Part 1: Messaging Framework (Second Edition). W3C Recommendation, April 2007.
- [30] Rosenberg, J., Schulzrinne, H. et al.: RFC 3261 - SIP: Session Initiation Protocol, IETF, June 2002.
- [31] Hedrick, C.: RFC 1058 - Routing Information Protocol, IETF, June 1988.
- [32] Moy, J.: RFC 2328 - OSPF Version 2, IETF, April 1998.
- [33] Coltun, R., Ferguson, D., Moy, J., Lindem, A.: RFC 5340 - OSPF for IPv6, IETF, July 2008.
- [34] ISO/IEC 10589:2002 Information technology -- Telecommunications and information exchange between systems -- Intermediate System to Intermediate System intra-domain routing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473). Second edition. ISO standard. November 2002, p.201.

- [35] Rekhter, Y., Li, T., Hares, S., A.: RFC 4271 - Border Gateway Protocol 4 (BGP-4), IETF, January 2006.
- [36] Schulzrinne, H., Casner, S., Frederick, R., Jacobson, V.: RFC 3550 - RTP: A Transport Protocol for Real-Time Applications, IETF, July 2003.
- [37] Huitema, C.: RFC 3605 - Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP), IETF, October 2003.
- [38] Schulzrinne, H., Rao, A., Lanphier, R.: RFC 2326 - Real Time Streaming Protocol (RTSP), IETF, April 1998.
- [39] Case, J., Fedor, M., Schoffstall, M., Davin, J.: RFC 1067 - Simple Network Management Protocol, IETF, August 1988.
- [40] McCloghrie, K., Rose, M., T.: RFC 1213 - Management Information Base for Network Management of TCP/IP-based internets: MIB-II, IETF, March 1991.
- [41] Case, J., D., McCloghrie, K., Rose, M., T., Waldbusser, S.: RFC 1901 - Introduction to Community-based SNMPv2, IETF, January 1996.
- [42] Harrington, D., Presuhn, R., Wijnen, B.: RFC 3411 - An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks, IETF, December 2002.
- [43] Enns, R., Bjorklund, M., Schoenwaelder, J., Bierman, A.: RFC 6241 - Network Configuration Protocol (NETCONF), IETF, June 2011.
- [44] EdgeCast Eyes CDN Federation,  
[http://www.lightreading.com/document.asp?doc\\_id=216113](http://www.lightreading.com/document.asp?doc_id=216113)
- [45] Level 3 Completes Substantial Expansion of Global CDN Capacity to Support Rising CDN Demand, PR Newswire, 2012, <https://www.prnewswire.com/news-releases/level-3-completes-substantial-expansion-of-global-cdn-capacity-to-support-rising-cdn-demand-147551265.html>
- [46] Draft-previdi-cdni-footprint-advertisement, IETF/CDNI WG,2012  
<http://www.potaroo.net/ietf/idref/draft-previdi-cdni-footprint-advertisement/>
- [47] MediaCloud Connect for Telcos, ISPs and MSOs, <http://www.mediamelon.com/isp-mediacloud.html>
- [48] Puopolo, S., Latouche, M., Le Faucheur, F., Defour, J.: Content Delivery Network (CDN) Federations, Cisco paper, 2011,  
[https://www.cisco.com/c/dam/en\\_us/about/ac79/docs/sp/CDN-PoV\\_IBSG.pdf](https://www.cisco.com/c/dam/en_us/about/ac79/docs/sp/CDN-PoV_IBSG.pdf)
- [49] U.S. National Institute of Standards and Technology (NIST) The NIST Definition of Cloud Computing, Sept. 2011, <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>



- [50] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I., Zaharia, M.: Above the Clouds: A Berkeley View of Cloud Computing, University of California at Berkeley.  
<https://www2.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>
- [51] Landis, C., Blacharski, D.: Cloud Computing Made Easy: An Easy to Understand Reference About Cloud Computing, Virtual Global Inc, 2013, ISBN 978-1482779424
- [52] Parkhill, D. F.: The Challenge of the Computer Utility, Addison-Wesley Publishing Company, 1966, ASIN: B000O121OS, ISBN: 0240507177
- [53] Regalado, A.: Who coined "Cloud Computing"?, MIT Technology Review, USA, Oct. 2011
- [54] Sadiku, M. N. O., Musa, S. M., Momoh, O.D.: Cloud computing: Opportunities and challenges, IEEE Potentials 02/2014; 33(1):34-36
- [55] Varia, J.: Architecting for the Cloud: Best Practices, January 2010
- [56] Mell, P., Grance, T.: Effectively and Securely Using the Cloud Computing Paradigm, NIST, 2009
- [57] Gorelik, E.: Cloud Computing Models, MIT, January 2013
- [58] The cloud tutorial, <http://thecloudtutorial.com>
- [59] Catteddu, D., Hogben, G.: Cloud Computing: Benefits, Risks and Recommendations for Information Security, ENISA, 2009;  
[www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment/at\\_download/fullReport](http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment/at_download/fullReport)
- [60] Cloud Security Alliance: Security Guidance for Critical Areas of Focus in Cloud Computing V2.1, <http://www.cloudsecurityalliance.org/csaguide.pdf>
- [61] OpenFlow Switch Specification. Version 1.5.1. Open Networking Foundation. March 26, 2015. p. 283
- [62] Network Functions Virtualisation - Introductory White Paper. SDN and OpenFlow World Congress, Darmstadt-Germany, October 22-24, 2012. p.16.
- [63] ETSI GS NFV-INF 003 V1.1.1 (2014-12), Network Functions Virtualisation (NFV); Infrastructure; Compute Domain. ETSI, December 2014
- [64] ETSI GS NFV-INF 004 V1.1.1 (2015-01), Network Functions Virtualisation (NFV); Infrastructure; Hypervisor Domain. ETSI, January 2015
- [65] ETSI GS NFV-INF 001 V1.1.1 (2015-01), Network Functions Virtualisation (NFV); Infrastructure Overview. ETSI, January 2015

- [66] ETSI GR NFV 001 V1.2.1 (2017-05), Network Functions Virtualisation (NFV); Use Cases. ETSI, May 2017
- [67] ETSI GS NFV-MAN 001 V1.1.1 (2014-12), Network Functions Virtualisation (NFV); Management and Orchestration. ETSI, December 2014
- [68] Guo, P. A Survey of Software as a Service Delivery Paradigm. TKK T-110.5190 Seminar on Internetworking, 2009
- [69] Kaysen, M. Understand the "SVOD", "TVOD" and "AVOD" terms and business models of streaming services like Netflix. 2015.  
<https://www.linkedin.com/pulse/understand-svod-tvod-avod-terms-business-models-streaming-mads-kaysen>
- [70] ITU. ZDF - HYBRID BROADCAST BROADBAND TELEVISION (HbbTV). Document WP 6B/[ZDF], 2012
- [71] HbbTV Forum Nederland. Overview of Interactive Television services according to the HbbTV standard in Europe. 2014. [http://hbbtv.nu/wp-content/uploads/2014/05/HbbTV\\_in\\_Europe\\_v5b\\_English.pdf](http://hbbtv.nu/wp-content/uploads/2014/05/HbbTV_in_Europe_v5b_English.pdf)
- [72] Chen, J., Yuan, L., Mingsins, C. Extending the Definition of E-Services and Its Implications to E-Services Development. International Joint Conference on Service Sciences, 2012, pp. 211-216
- [73] Mehdi K.-P. Encyclopedia of E-Commerce, E-Government, and Mobile Commerce. Idea Group Inc., 2006. p. 1260. ISBN 1-59140-799-0
- [74] Tarmo, K. and Ain, A. The Development of eServices in an Enlarged EU: eGovernment and eHealth in Estonia. EC JRC Technical Report, 2008. ISSN 1018-5593
- [75] Weber, R. H., Internet of Things - New Security and Privacy Challenges, Computer Law & Security Review 26: 23-30, 2010
- [76] Podhradský, P., Mikóczy, E., Lábaj, O., Londák, J., Trúchly, P., at al: NGN Architectures and NGN Protocols, LdV IntEleCT, Educational publication, 210 pages, Published by ČVUT Praha, ISBN: ISBN:978-80-01-04949-5, September 2011
- [77] ITU-T Recommendation Y.1910 (09/2008), IPTV functional architecture, ITU-T, 2008
- [78] Mikóczy, E. Advanced Multimedia Architecture for Next Generation of Internet Protocol Television Systems. Dissertation theses, FEI STU Bratislava, 2010
- [79] W3C. WebRTC 1.0: Real-time Communication between Browsers. W3C Editor's Draft 22 December 2015. <http://w3c.github.io/webrtc-pc/>
- [80] WebRTC homepage. <https://webrtc.org/>

- [81] Rao, U. H., Nayak, U., The InfoSec Handbook - An Introduction to Information Security, Apress Media, New York, 2014, ISBN 978-1-4302-6382-1
- [82] Moghaddasi, H., Sajjadi, S., Kamkarhaghghi, M., Reasons in Support of Data Security and Data Security Management as Two Independent Concepts: A New Model, The Open Medical Informatics J., Vol. 10, 2016, pp 4-10
- [83] ITU Recommendation X.800 - Data communication networks: Open systems interconnection (OSI) Security, structure and applications - Security architecture for open systems interconnection for CCITT applications, Geneva, 1991
- [84] Stallings, W., Cryptography and network security - Principles and practice, 5th ed., Pearson Education, New York, 2011
- [85] Ruff, A., Network Security 1 and 2 Companion Guide (Cisco Networking Academy), Cisco Press, 2006, ISBN: 978-1587131622
- [86] JPEG, ITU recommendation T.81, 09/1992
- [87] Ghanbari, M., Standard Codecs: Image Compression to Advanced Video Coding, Institution of Electrical Engineers, 2003, 407 pages, ISBN:0852967101
- [88] JPEG 2000 standard, ISO/IEC 15444-1, 15444-2
- [89] MPEG-2, ISO/IEC 13818 standards, ISO/IEC 13818-1 to ISO/IEC 13818-11
- [90] MPEG-4 AVC standard, ISO/IEC 14496-10, ITU-T H.264
- [91] Schwarz, H., Marpe, D., Wiegand, T., Overview of the scalable video coding extension of the H.264/AVC standard, IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no.9, pp 1103-1120, 2007
- [92] MPEG-H standard part 2, ISO/IEC 23008
- [93] MPEG-21 Digital Item adaptation, ISO/IEC 21000-7:2004 - Part 7: Digital Item Adaptation
- [94] Sodagar, I., The MPEG-DASH Standard for Multimedia Streaming Over the Internet. IEEE MultiMedia 18, 4 (October 2011), 62-67.  
DOI=<http://dx.doi.org/10.1109/MMUL.2011.71>
- [95] Sieber, C., et al., Implementation and User-centric Comparison of a Novel Adaptation Logic for DASH with SVC, Proc. IFIP/IEEE International Symposium on Integrated Network Management, 2013, pp. 1318-1323.
- [96] Examples of DASH SVC datasets, 2017, <http://concert.itec.aau.at/SVCDataset/>
- [97] Mueller, Ch., MPEG-DASH (Dynamic Adaptive Streaming over HTTP, ISO/IEC 23009-1), <https://bitmovin.com/dynamic-adaptive-streaming-http-mpeg-dash/>

- [98] Online Learning Adaptation Strategy for DASH Clients, MMSys 2016 Klagenfurt, Austria, 2016, ACM, ISBN 978-1-4503-2138-9, DOI: 10.1145/1235
- [99] Xylomenos, G., Ververidis, C. N., Siris, V. A., Fotiou, N., Tsilopoulos, C., Vasilakos, X., Katsaros, K. V., Polyzos, G. C., A survey of information-centric networking research. *IEEE Communications Surveys Tutorials*, 16(2):1024-1049, 2014
- [100] Ge, C., Wang, N., Skillman, S., Foster, G., Cao, Y., QoE-Driven DASH Video Caching and Adaptation at 5G Mobile Edge. In *Proceedings of the 3rd ACM Conference on Information-Centric Networking (ACM-ICN '16)*. ACM, New York, USA, 2016, 237-242
- [101] Fajardo, J. O., Taboada, I., Liberal, F., Improving content delivery efficiency through multi-layer mobile edge adaptation, *IEEE Network*, 29(6):40-46, Nov. 2015
- [102] ETSI White Paper No. 11, Mobile Edge Computing A key technology towards 5G, [http://www.etsi.org/images/files/ETSIWhitePapers/etsi\\_wp11\\_mec\\_a\\_key\\_technology\\_towards\\_5g.pdf](http://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp11_mec_a_key_technology_towards_5g.pdf)
- [103] Ye, Y., He, Y., Wang, Y. k., Hendry, SHVC, the Scalable Extensions of HEVC, and Its Applications, *ZTE COMMUNICATIONS* February 2016 Vol.14 No.1, <http://www.cnki.net/kcms/detail/34.1294.TN.20160122.1848.004.html>
- [104] Weil, N., The State of MPEG-DASH 2017, <http://www.streamingmediaglobal.com/Articles/Editorial/Featured-Articles/The-State-of-MPEG-DASH-2017-116505.aspx>
- [105] Skupin, R., Sanchez, Y., Podborski, D., Hellge, C., Schierl, T., HEVC tile based streaming to head mounted displays, *Consumer Communications & Networking Conference (CCNC)*, 2017 14th IEEE Annual, DOI: 10.1109/CCNC.2017.7983191
- [106] Recommendation ITU-R M.2083-0 (2015-09), *IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond*. ITU-R, September 2015
- [107] 3GPP Release 15: <http://www.3gpp.org/release-15>
- [108] 3GPP Release 16 <http://www.3gpp.org/release-16>
- [109] DRAFT NEW REPORT ITU-R M.[IMT-2020.TECH PERF REQ], Minimum requirements related to technical performance for IMT-2020 radio interface(s). ITU-R, February 2017
- [110] 3GPP TS 22.261 V16.2.0 (2017-12), *Service requirements for the 5G system; Stage 1 (Release 16)*, 3GPP, December 2017
- [111] 3GPP TS 23.501 V15.0.0 (2017-12), *System Architecture for the 5G System; Stage 2 (Release 15)*, 3GPP, December 2017

- [112] 3GPP TS 23.502 V15.0.0 (2017-12), Procedures for the 5G System; Stage 2, (Release 15), 3GPP, December 2017
- [113] 3GPP TS 23.503 V15.0.0 (2017-12), Policy and Charging Control Framework for the 5G System; Stage 2 (Release 15), 3GPP, December 2017
- [114] 3GPP TS 29.060 V15.1.0 (2017-12), GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface (Release 15). 3GPP, December 2017
- [115] ETSI GS NFV-MAN 001 V1.1.1 (2014-12) Network Functions Virtualisation (NFV); Management and Orchestration. ETSI, December 2014
- [116] ETSI TS 123 101 V14.0.0 (2017-05), General Universal Mobile Telecommunications System (UMTS) architecture. ETSI, May 2017
- [117] 5G PPP Architecture Working Group, View on 5G Architecture (Version 2.0).  
[https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017\\_For-Public-Consultation.pdf](https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017_For-Public-Consultation.pdf)
- [118] 3GPP TS 22.071 V14.1.0 (2015-09), IP Multimedia Subsystem (IMS); Stage 2 (Release 15), 3GPP, December 2017
- [119] 3GPP TS 22.268 V15.1.0 (2017-06), Public Warning System (PWS) requirements; (Release 14), 3GPP, June 2017
- [120] 3GPP TS 23.228 V15.1.0 (2017-12), Location Services (LCS); Stage 1 (Release 14), 3GPP, September 2017
- [121] 5G, the Internet of Things (IoT) and Wearable Devices. GSMA.  
[https://www.gsma.com/publicpolicy/wp-content/uploads/2017/10/5g\\_iot\\_web\\_FINAL.pdf](https://www.gsma.com/publicpolicy/wp-content/uploads/2017/10/5g_iot_web_FINAL.pdf)
- [122] Huawei - SDN / NFV Innovation Workshop, Nir Halachmi, Network Research Product Management
- [123] Rendering Omni-directional Stereo Content. Google Inc.  
<https://developers.google.com/vr/jump/rendering-ods-content.pdf>
- [124] ITU-T: Overview of the Internet of things.  
[https://www.itu.int/rec/dologin\\_pub.asp?lang=e&id=T-REC-Y.2060-201206-I!!PDF-E&type=items](https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-Y.2060-201206-I!!PDF-E&type=items)
- [125] Internet of Vehicles - Technologies and Services Third International Conference, IOV 2016, Nadi, Fiji, December 7-10, 2016, Proceedings
- [126] A new kind of smart car. <https://newsroom.cisco.com/feature-content?articleId=1724437>

- [127] Smart Cities The importance of a smart ICT infrastructure for smart cities.  
<https://www.stokab.se/Documents/Nyheter%20bilagor/SmartCityInfraEn.pdf>
- [128] Deploying 5G-Technologies in Smart City and Smart Home Wireless Sensor Networks with Interferences. <https://rd.springer.com/article/10.1007/s11277-015-2480-5>
- [129] Telia delivers smart mailboxes to Finnish postal service.  
<https://internetofbusiness.com/telia-makes-postboxes-smarter-finnish-posti/>
- [130] Smart Wearables in Healthcare: Signal Processing, Device Development, and Clinical Applications. <https://www.hindawi.com/journals/jhe/si/389732/cfp/>
- [131] 5G Cellular Networks Are the Future of Robotics.  
<https://eu.mouser.com/applications/robotics-and-5g/>
- [132] 5G network slicing enables safer “eye in the sky” for drones.  
<http://www.telecomtv.com/articles/5g/5g-network-slicing-enables-safer-eye-in-the-sky-for-drones-15873/>
- [133] Applications and use cases of 5G for IoT solutions, AR and VR.  
<https://medium.com/@patelnisha121/applications-and-use-cases-of-5g-for-iot-solutions-ar-and-vr-37072a2a1711>