Building network intelligence – Traffic classification

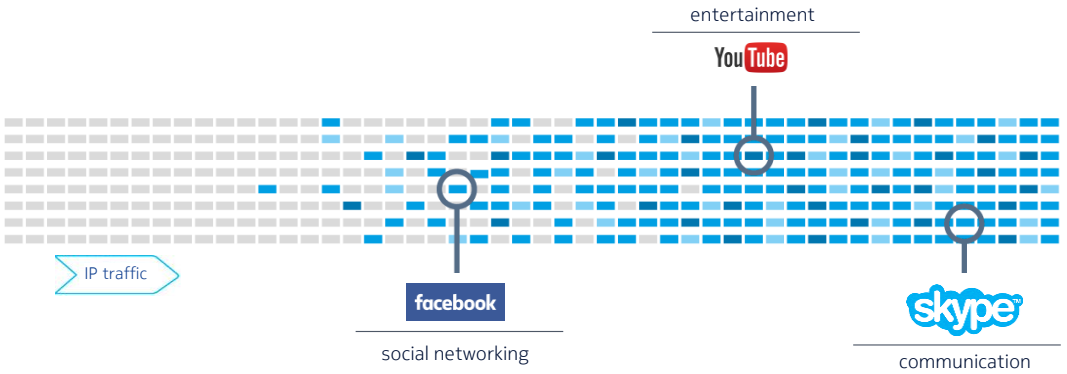
- Marios Bougioukos, Ph.D.

16-01-2017

What Traffic Classification stands for?

A key role for Intelligent Networks and Network Management

- Is the identification and categorization of the traffic mixture present in IP-based networks



- Helps to understand the network applications behavior
- Enables network service-awareness
- Provides network with intelligence

Intelligence

- optimized resource utilization
- bandwidth management
- better usage of existing infrastructure

Management

NOKIA

Traffic Classification

Capacity and benefits

- Service-differentiation
  - each type of service can be handled in a more “personal” way
  - Quality of Service (QoS) improvement
  - Quality of Experience (QoE) improvement

- Advertising
  - statistics for marketing purposes
  - data analytics
- Security
  - malicious traffic detection
  - prohibited sites or contents
- Accounting
  - charging differentiation
  - user specific charging policy

- Network design and engineering
  - better bandwidth management
  - optimization by tuning some network parameters
  - offloading some unimportant application traffic
- Research
  - to imitate the real traffic of applications in the network
  - study and understand the behavior of different applications

Mobile data traffic

Dominant in the Global Internet traffic

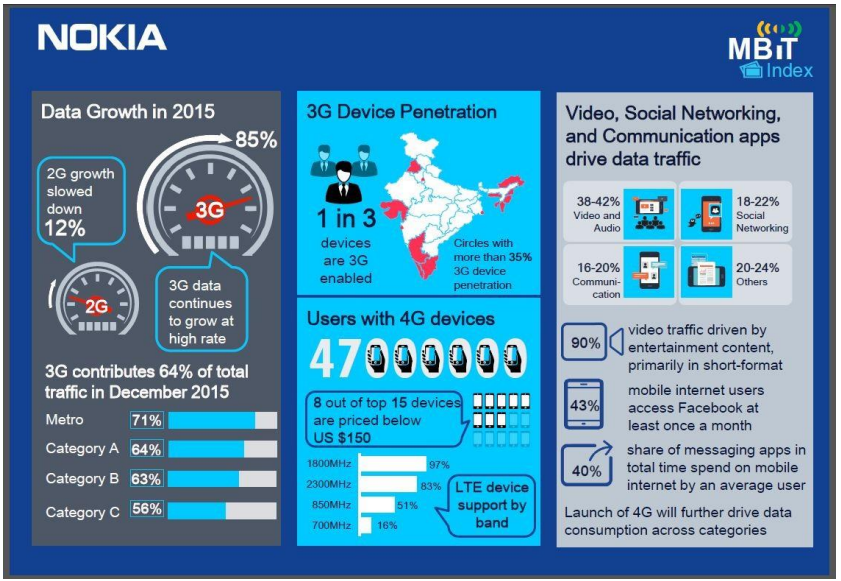


Figure:

- 3G data continued to grow at high rate (85% YoY) and now contributes more than 50% of data traffic across categories of circles
- Mobile traffic is driven by consumption of Video, Social Networking, and Communication related content, all of which combined contribute 90% of mobile traffic

should focus on the Mobile Internet traffic, which will dominate the Global Internet traffic

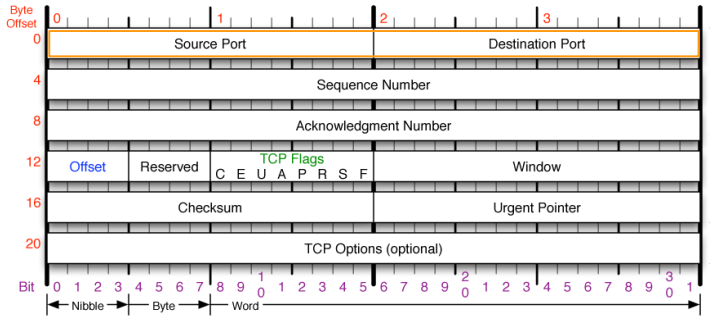
Traffic Classification approaches (1)

Port-based method

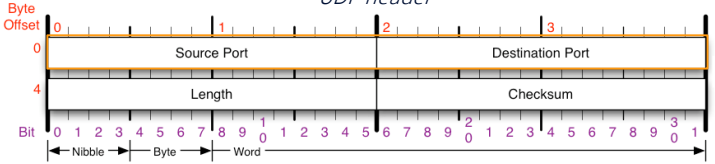
- The first and the simplest one
- Examines the IP packet in the transport layer (Layer-4) for well-known source/destination ports as these assigned to the IANA

Assigned Port	Application
20	FTP Data
21	FTP Control
22	SSH
23	Telnet
25	SMTP
53	DNS
80	HTTP
110	POP3
123	NTP
161	SNMP
3724	WoW

TCP header



UDP header



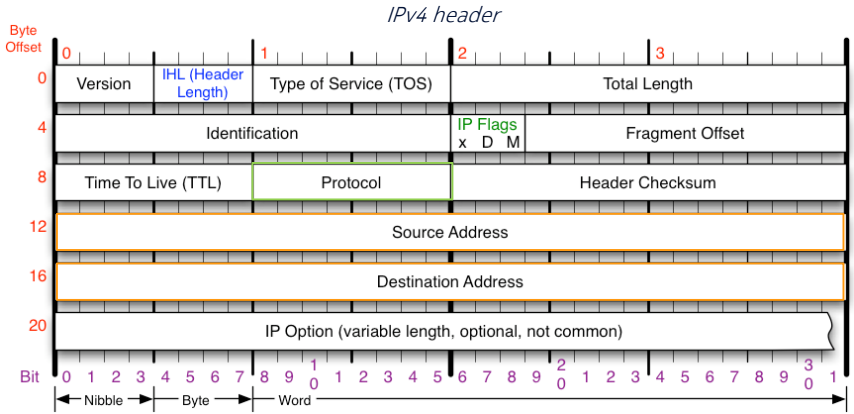
- Other transport protocols: SCTP and DCCP

- Problems exist for dynamic ports or abuse of well-known ports
- Only basic applications (client-to-server-like) can be identified

Traffic Classification approaches (2)

IP-based method

- Follows the same approach as the port-based technique
- Examines the IP packet in the network layer (Layer-3) for well-known source/destination IP addresses
- Periodic updates should be applied in order to maintain this technique accurate



- A combination of the port-based and IP-based method is the service-based method
- A service is defined as a triplet <IP, Port, Protocol> assigned to a specific application, ex. Skype <134.170.16.141, 2007, 17>
- A database of all these services is created in an off-line phase using a dataset of labeled flows and is used in turn for real-time classification

Traffic Classification approaches (3)

Payload-based method

- Inspection of the packet payload (Layer-7) searching for unique application signatures (pattern matching)
- This method also called Deep Packet Inspection (DPI) technique
- Most effort focuses on the identification of P2P applications because they use camouflage strategies (abuse of well-known ports)

Patterns extracted from well-known P2P applications

P2P Protocol	*String*	*Trans. prot.*	*Def. ports*
eDonkey2000	0xe319010000	TCP/UDP	4661-4665
	0xc53f010000		
Fasttrack	"Get /.hash"	TCP	1214
	0x270000002980	UDP	
BitTorrent	"0x13Bit"	TCP	6881-6889
Gnutella	"GNUT", "GIV"	TCP	6346-6347
	"GND"	UDP	
MP2P	GO!!, MD5, SIZ0x20	TCP	41170 UDP
Direct Connect	"$MyN","$Dir"	TCP	411-412
	"$SR"	UDP	
Ares	"GET hash:"	TCP	-
	"Get sha1:"		

* *Karagiannis et al.*

Advantages:

1. High accurate
2. Successful with application using random ports (P2P) and tunneled traffic

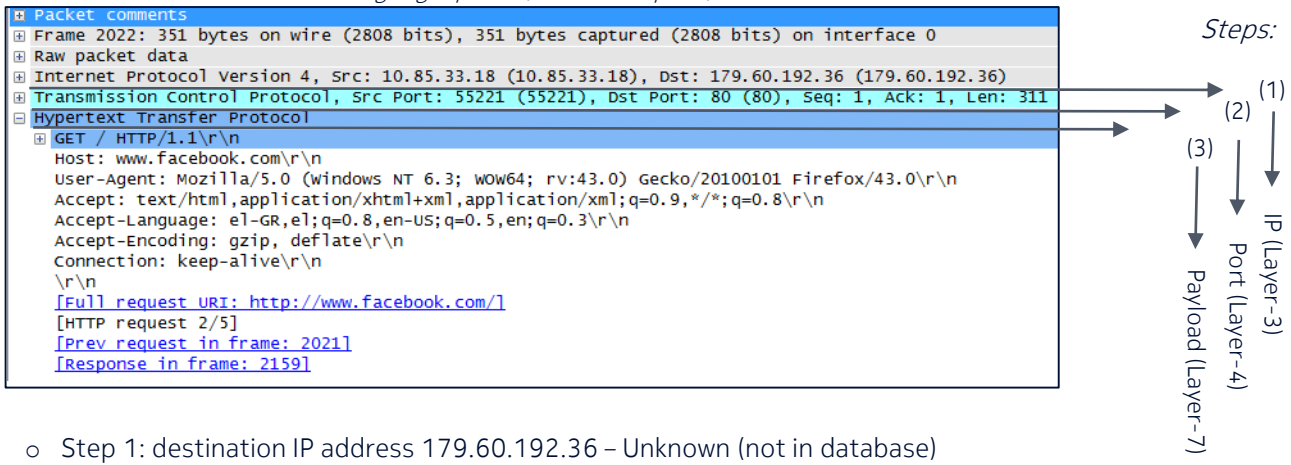
Disadvantages:

1. High resource requirements for the pattern searching
2. Encrypted traffic limitation
3. Continuous update (signatures must be kept up-to-date)

Traffic Classification approaches (4)

A Traffic Classification example

Outgoing IP packet (Wireshark capture)



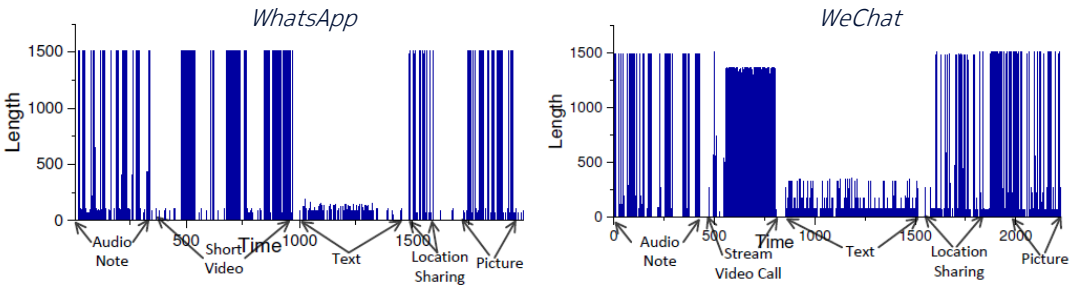
- Step 1: destination IP address 179.60.192.36 – Unknown (not in database)
  - Not satisfied – I want to find the protocol
  - Let's go deeper!
- Step 2: destination port 80 - HTTP application (Web access)
  - Not satisfied – I want to find the host
  - Let's go deeper!
- Step 3: host-name www.facebook.com - Facebook site (create metadata)
  - Yes!

Traffic Classification approaches (5)

Statistic-based method

- Uses statistical characteristics from network traffic such as:
  - distribution of packet size
  - packet inter-arrival time
  - number of traffic
  - traffic rate

- Advantage of identifying traffic type without packet inspection. Encrypted data are also covered



**Yanjie Fu et al.*

- Service-type identification instead of specific mobile application type

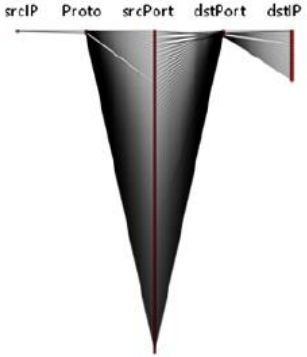
- Network conditions and high bit error rates may affect the accuracy of the classification results

Traffic Classification approaches (6)

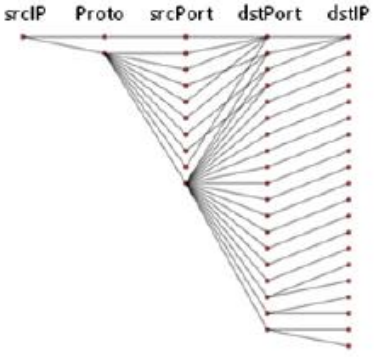
Host behavior-based method

- Uses communication patterns or causality of application traffic extracted from network and transport layers
- Identification of mobile applications and certain traffic categories (Web, P2P) since they use constant and unique characteristics

Web



Skype (voice)



**Mongkolluksamee et al.*

Graphlets:

Web: Client-Server communication pattern
Protocols: TCP, UDP (DNS messages)
Destination ports: 80, 443, 53

Skype (voice): P2P communication pattern
Protocols: TCP, UDP

- Requires periodic updates to adapt the classification model to new applications

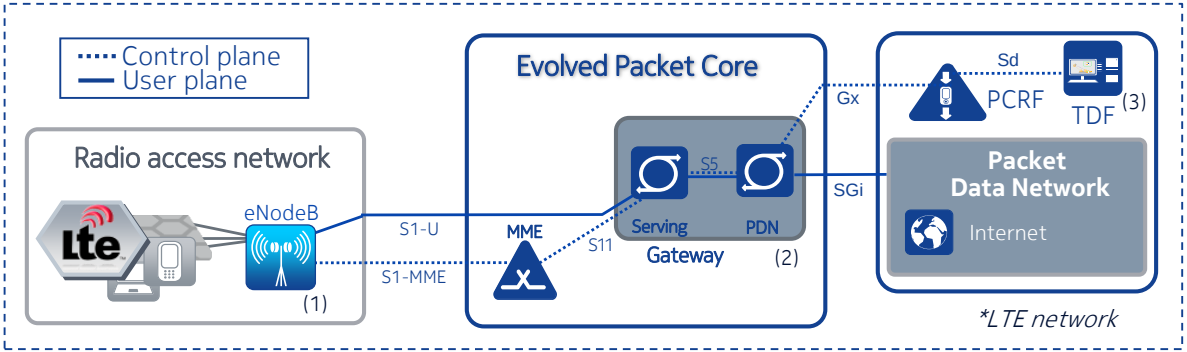
Deep Packet Inspection Tools

Off-the-shelf solutions

- Deep Packet Inspection (DPI) is considered as the most accurate traffic classification method
- DPI actually includes all the aforementioned classification methods
- PACE (https://www.ipoque.com/products/dpi-engine-rsrpace-2)
  - classifies over 95% of network traffic
  - up to 9 Gbps per core
  - identification of 2800 application and protocols
  - Bitrate, latency, aspect ratio and other approaches
- Qosmos (http://www.qosmos.com/)
  - recognizes over 97% of network traffic
  - covers over 2000 protocols and applications
  - real-time traffic analysis at 10, 40 or 100 Gbps
  - methods include application metadata, behavioral and statistical analysis, etc.
- nDPI (http://www.ntop.org/)
  - supports more than 100 protocols and applications
  - analysis of session certificates
  - open source

Locating DPI functionality in Mobile Networks

Modules performing traffic analysis and policy



(1)

eNodeB

- DSCP marking

(2)

Packet Data Network Gateway

- policy
- charging
- firewall

(3)

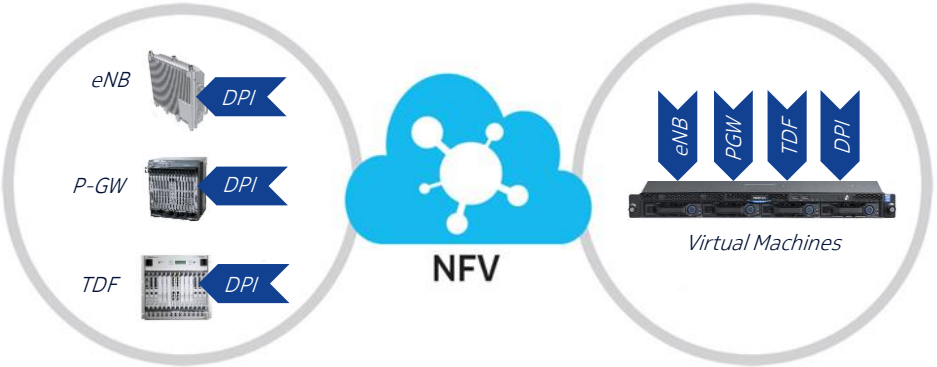
Traffic Detection Function (TDF)

- gating actions (drop, pass)
- traffic shaping
- redirection

NOKIA

Realizing DPI functionality in the SDN world (1)

The advantage of NFV technology



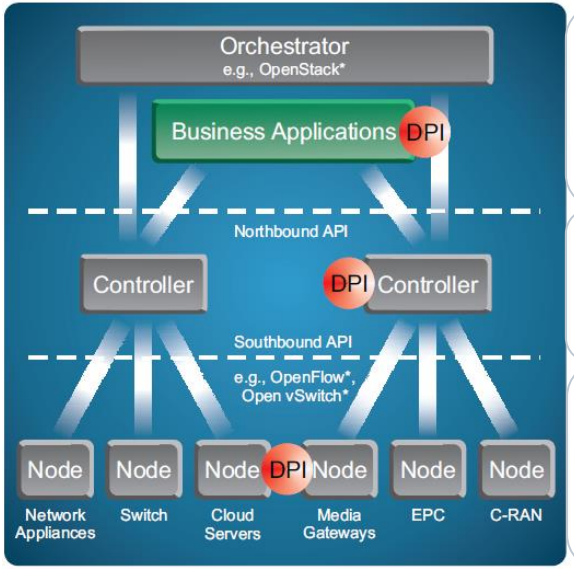
- DPI is implemented in several and different network devices
- High cost to implement DPI technology many times
- Different classification methods may lead to different results

- With SDN and NFV, DPI can become a shared function
- Low cost since DPI is implemented in fewer devices
- Easier application interconnection, easy to apply a consistent format for metadata

Realizing DPI functionality in the SDN world (2)

Locating DPI in the SDN layers

- With SDN approach DPI information and therefore network intelligence is shared through the network
- Lowers CPU and energy costs as application recognition is performed once



Application layer

A small traffic percentage is transmitted from nodes to the applications (via controller) running DPI

Best performance for non-time-critical applications

Control layer

Intelligence is used for the control services or the upper layer

Duplicate packets are transmitted from nodes to the controller for L7 analysis (DPI)

Infrastructure layer

Nodes after identification can either apply a pre-defined policy or send information to the SND controller or the applications

Low latency, but costly

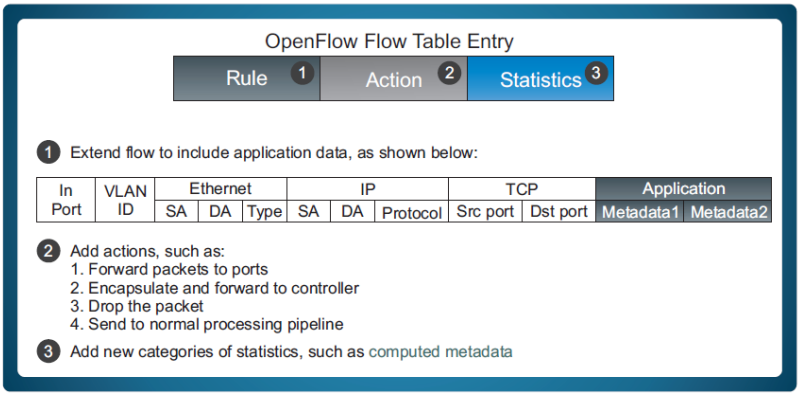
** Intel*

Realizing DPI functionality in the SDN world (3)

L7 information using OpenFlow protocol

- Currently, SDN has L2-4 visibility. For computed metadata per flow between nodes and SDN controller new L7 fields are required

- OpenFlow protocol must be enhanced to support L7 intelligence

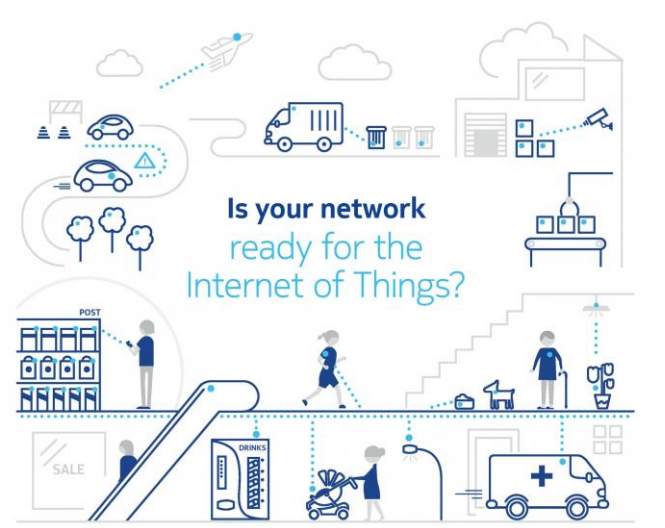


** Market Education Committee (MEC) of Open Networking Foundation (ONF)*

- Rules: identifications of protocol, applications (App ID), and metadata
- Actions: such as drop the packet, encapsulate and forward the packet to the controller, of forward packet to specific port
- Statistics: including computed metadata, HTTP-host name, HTTP cookie, and vendor-specific attributes (VSA)

In the IoT world

The role of Traffic Classification

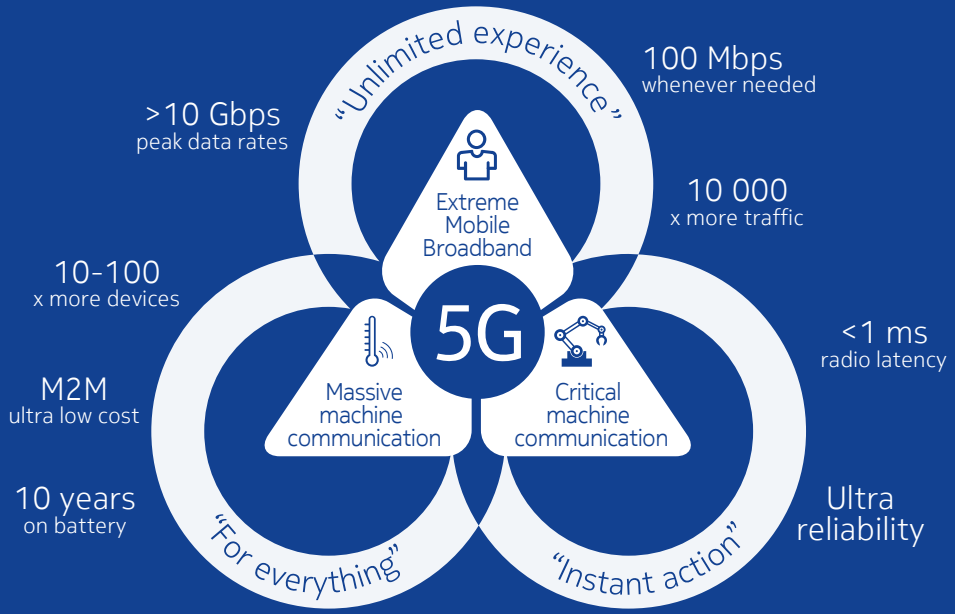


- IoT environment accommodates complex traffic due to sophisticated applications
- New or enhanced data protocols will be emerged (ex: MQTT, CoAP, AMQP, Websocket, Node)
- More personal devices connected, more personal data possible available to non-authorized users
- Several services require special handling (ex eHealth tools)

Real-time traffic classification is the wheel for a safe, qualitative and well-functioning IP-based future

5G era

Heterogeneous use cases – diverse requirements



  Public

More Information

In our web sites

- Latest Launches

  https://networks.nokia.com/latest-launches

- Research Publications

  https://www.bell-labs.com/our-research/publications/

- Technology Vision

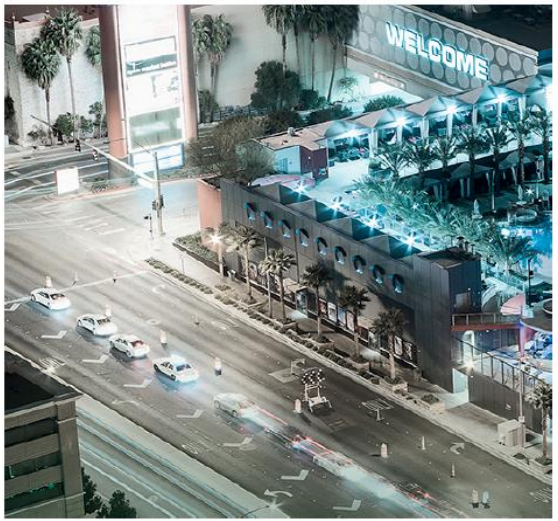
  http://networks.nokia.com/innovation/technology-vision

- 5G

  http://networks.nokia.com/innovation/5g

NOKIA

Thank you - Connect with us - Any Questions?



CONTACT

Location 1: Agisilaou 6-8, 15123 Marousi
Location 2: 14klm. National Road Athens-Lamia
Switchboard.: 210 6253008
Fax: 210 6206002
Web address:
http://careers.networks.nokia.com/
E-mail: tc.athens@nokia.com

SOCIAL MEDIA

https://www.facebook.com/nokiaglobalcareers
https://twitter.com/nokianetworks
https://linkedin.com/company/nokia

networks.nokia.com

NOKIA