# IP Networks Operation

stefano.gollinucci@vodafone.com

Torino, January 14th, 2021

# Agenda

1 Network overview: access network, legacy core network, NFV

2 IP backhauling, IP Core and underlying TX

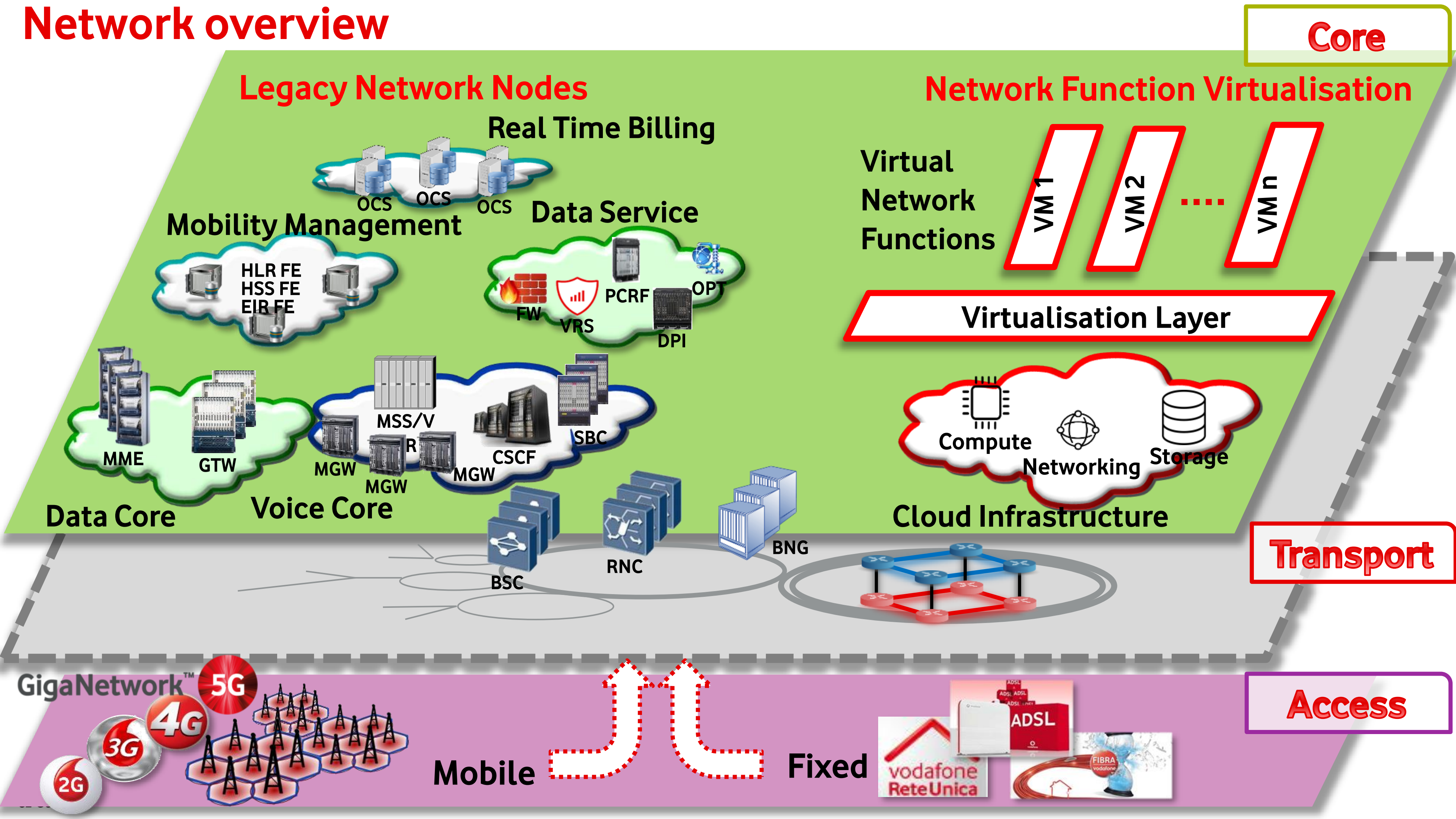
3 IP MPLS and Segment Routing

4 SDN

5 IP Network Operation Challenges & Evolution

6 Wrap-Up and Q&A

# Network overview



# Agenda

1 Network overview: access network, legacy core network, NFV

2 IP backhauling, IP Core and underlying TX

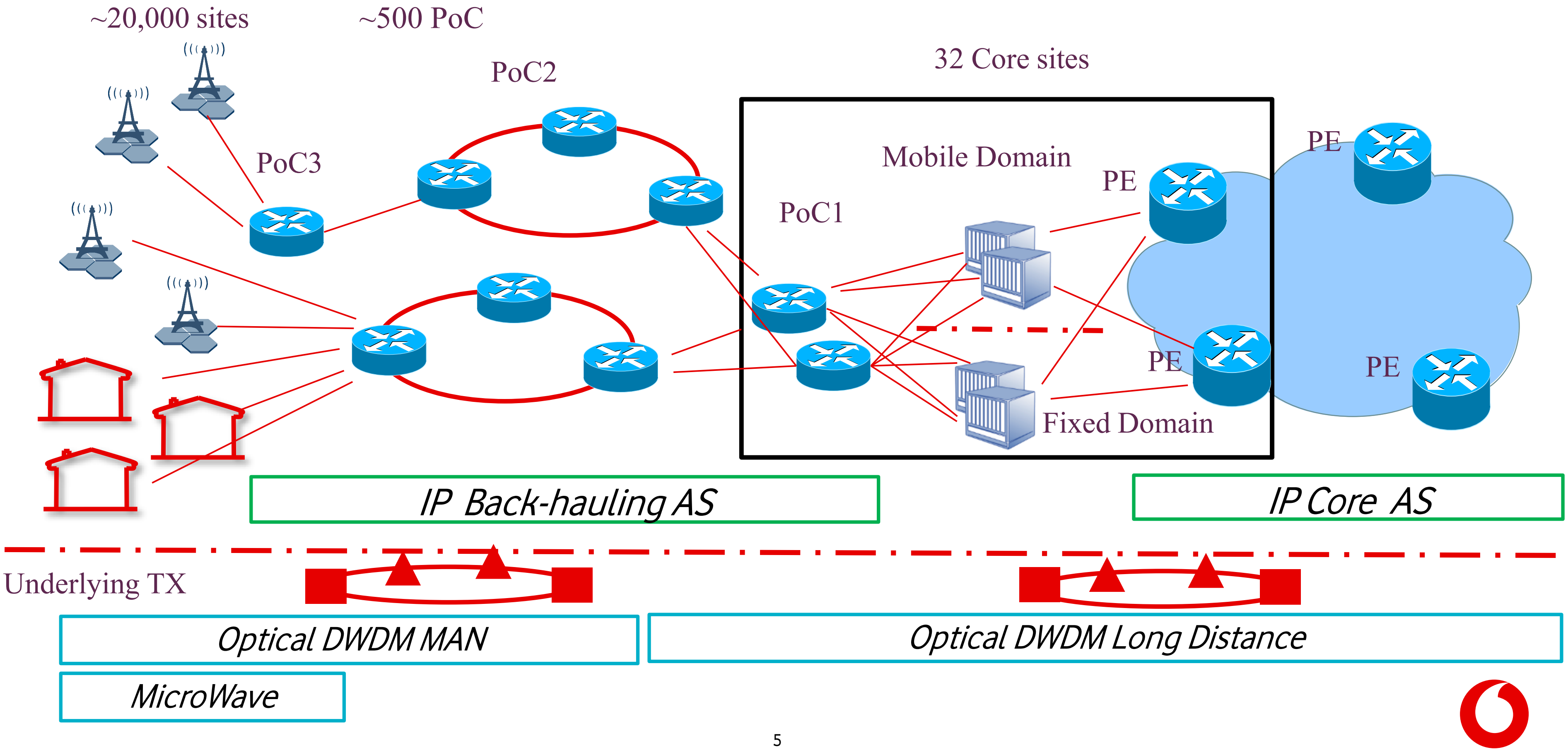
3 IP MPLS and Segment Routing

4 SDN

5 IP Network Operation Challenges & Evolution

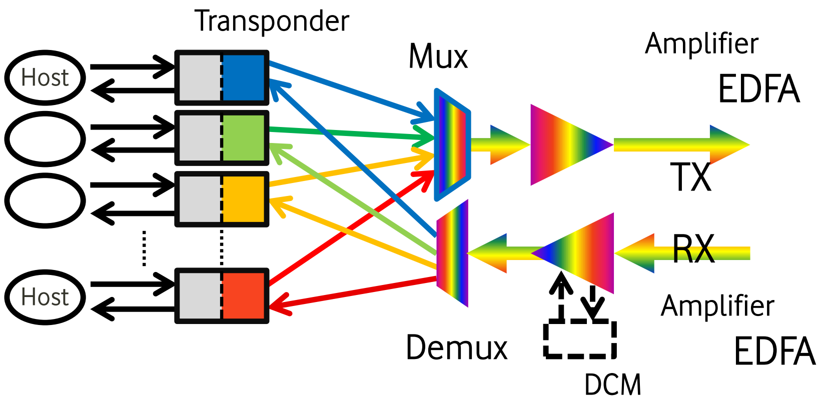
6 Wrap-Up and Q&A

# Transport network: mobile-fixed convergence on IP networks

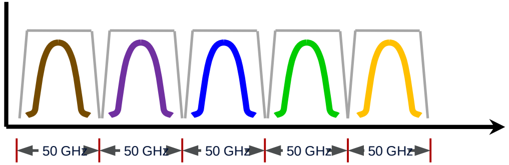


# The underlying TX layer: DWDM

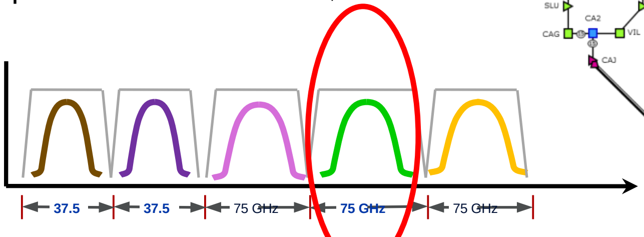
DWDM (Dense Wavelength Division Mux)
3rd window (1,550 nm)



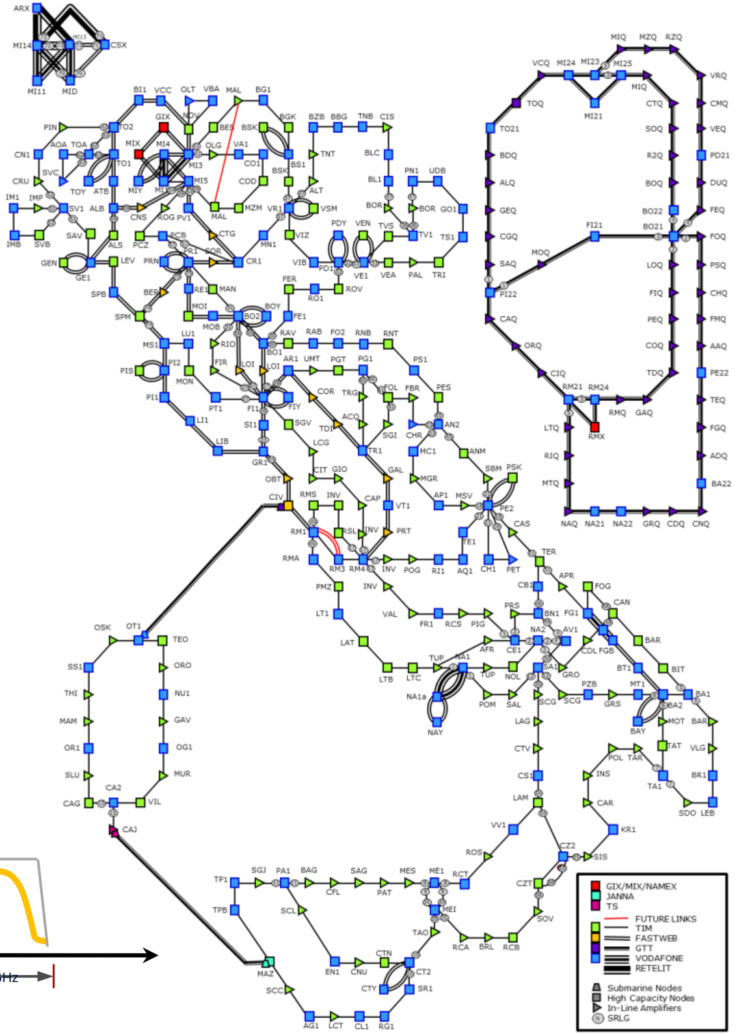
~240 ROADM nodes (Reconfigurable optical add-drop mux)
~140 regeneration nodes
FlexGrid technology with flexible channel size and frequency.
Standard 50Ghz and probabilistic shape modulation 75GHz,
Without fixed filter constraints



Fixed 50 GHz Grid

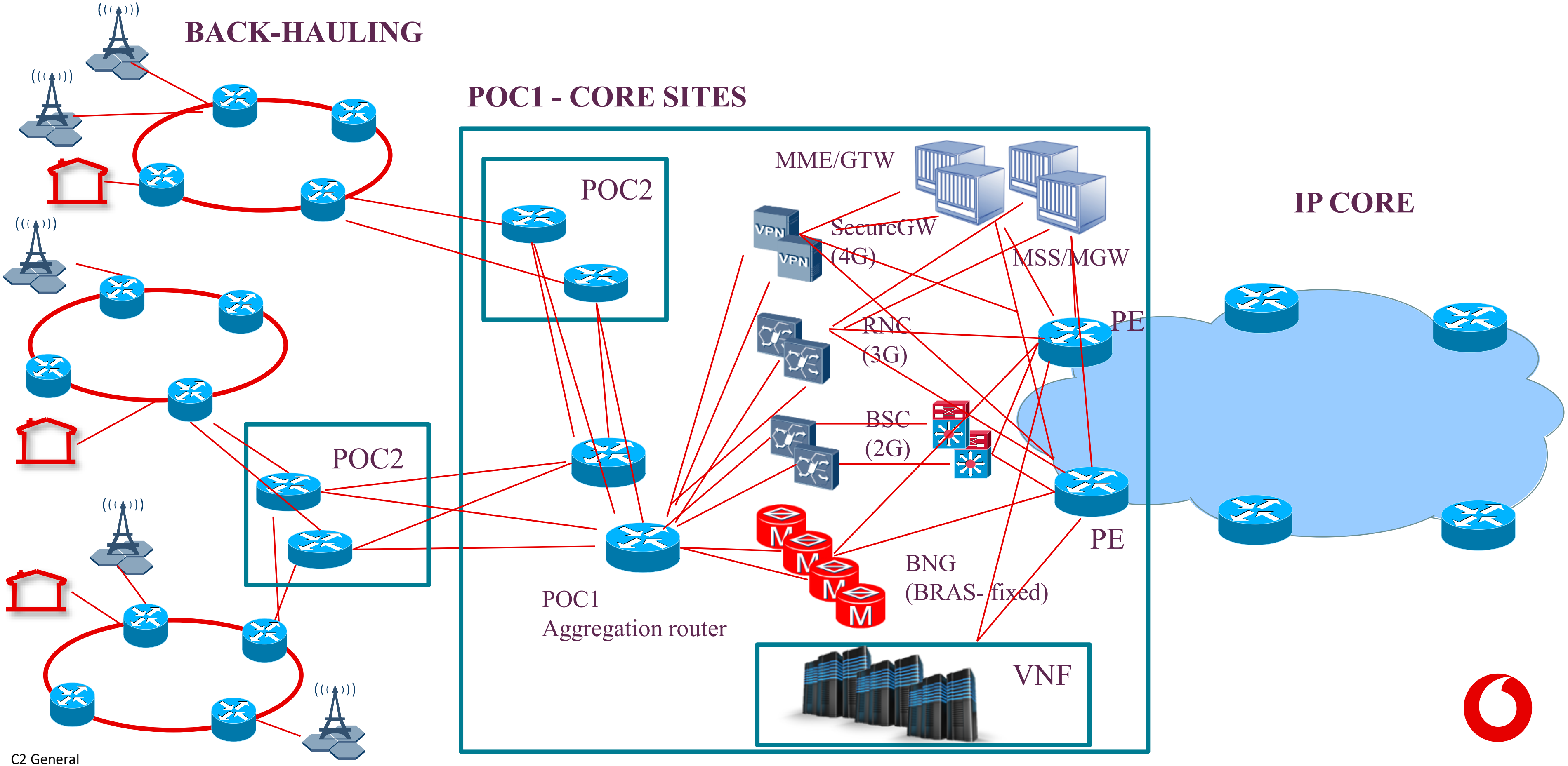


Flexible Grid



C2 General

# IP Core – Core sites overview



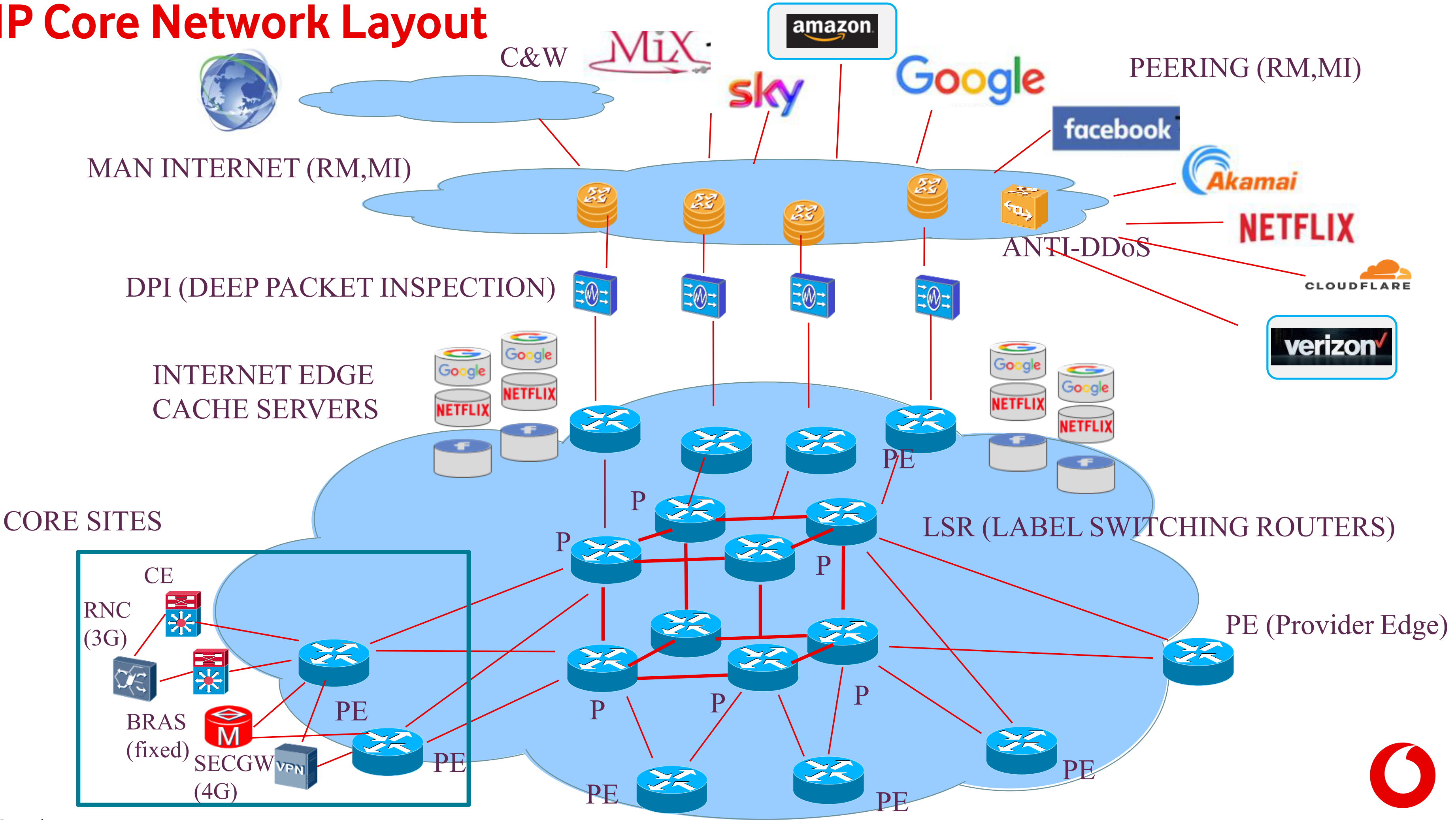
C2 General

# IP Core – POC1 Core sites

IP Core is deployed in 32 Core sites, each of them including:

- Redundant POC1 aggregation routers to collect traffic from IP backhauling POC2 network
- Legacy network platforms, among which:
  - BNG (BRAS) servers to manage fixed customers data traffic towards IP Core
  - 2G, 3G controllers to manage mobile traffic, they forward voice traffic towards Voice Core, and data traffic towards Data Core
  - Redundant 4G SecureGateways, terminating IPSEC tunneling, and forwarding VOLTE (Voice over 4G) towards IMS Voice Core and data traffic towards Data Core
  - Voice Core systems, MSS, MGW, IMS, etc.
  - Data Core systems, MME, GTW, etc.
- NFV datacenters, which provide shared resources (hardware, storage, computing) to Virtual Machines implementing the Network Functions above described
- Redundant PE routers as MPLS ingress points of IP core

# IP Core Network Layout



# IP Core Network

The IP Core consists of:

- 64 PE  routers (Label Switching Edge) located in 32 Core sites
- 8 P routers (Label Switching Router) in a two-layers fully redundant architecture
- 8 Route Reflectors to manage iBGP sessions scalability issue
- Dedicated PEs towards Internet, with cache servers to increase efficiency with OTTs traffic and to announce eBGP routes
- DPI (Deep Packet Inspection) layer to implement optimisation policies on Internet traffic
- Internet Peering points  as well as interconnections with internatonal carrier in Roma, Milano
- Anti-DDOS systems

# Agenda

1 Network overview: access network, legacy core network, NFV

2 IP backhauling, IP Core and underlying TX

3 IP MPLS and Segment Routing

4 SDN

5 IP Network Operation Challenges & Evolution

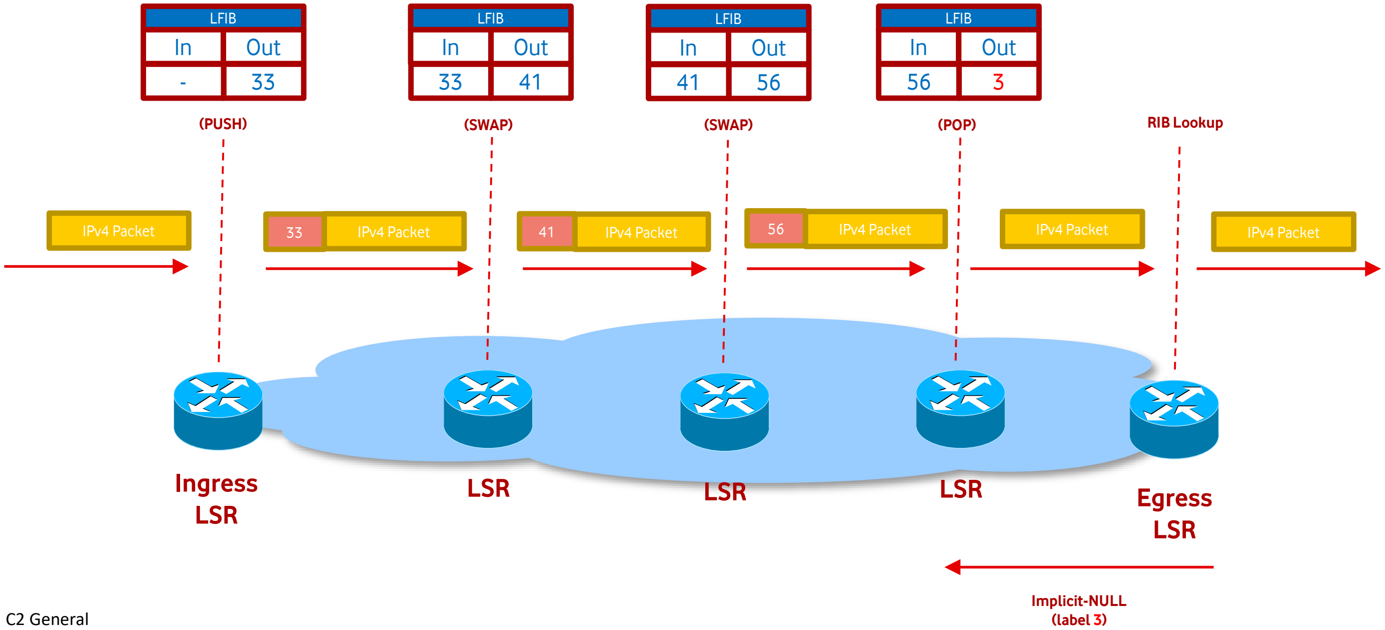
6 Wrap-Up and Q&A

# MPLS (Multi Protocol Label Switching)

- Different services with different service requirements (latency, bandwidth, reliability, etc.)
- MPLS makes it possible to segregate traffic flows through the creation of VPNs (Virtual Private Networks)
- MPLS implements FEC (Forwarding Equivalence Class), that is a group of IP packets which are forwarded in the same manner, over the same path, and with the same forwarding treatment. While in a plain IP network the FEC is determined at each hop, on an MPLS network the FEC is determined once, at the ingress of the network.
- Routing is based on distribution and swap of labels between routers rather than less efficient IP routing table lookup
- Traffic engineering is supported through the creation of MPLS tunnels or LSPs (Label Switched Paths)

# Label Switching

- At the ingress point Provider Edge (PE) routers "push" labels to IP packets of the specific traffic flow
- Intermediate Label Switch Routers (LSR or P routers), "swap" labels to select the path
- At the egress point PE routers "pop" the labels and and perform local Routing Information Base (RIB) lookup (Penultimate Hop Popping may be used to off-load PEs)
- Creation of VPN and traffic engineering are supported through L3VPN and RSVP-TE protocols respectively

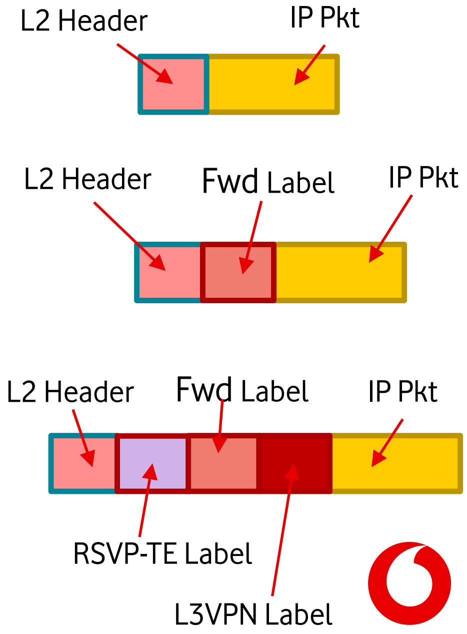


| LFIB | |
|---|---|
| In | Out |
| - | 33 |

| LFIB | |
|---|---|
| In | Out |
| 33 | 41 |

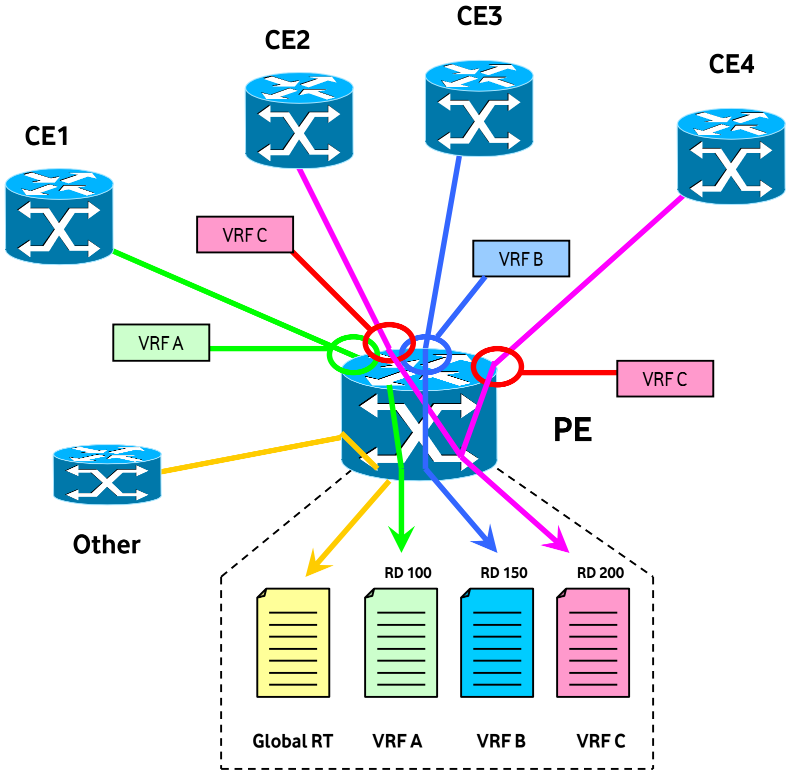
| LFIB | |
|---|---|
| In | Out |
| 41 | 56 |

| LFIB | |
|---|---|
| In | Out |
| 56 | 3 |



# VRF – VIRTUAL ROUTING AND FORWARDING

Provider Edge (PE) routers segregate traffic of different VPNs creating VRFs



MPLS PEs support the creation VRFs. Each **VRF** constitutes a separate routing and forwarding table, isolated from the others.

The "regular" routing table is called **Global Routing Table**, and routes/packets default to this table when a VRF is not specified

*VRF names have only local significance*. Having the same VRF name among different routers does not a mean the two VRFs are part of the same VPN.

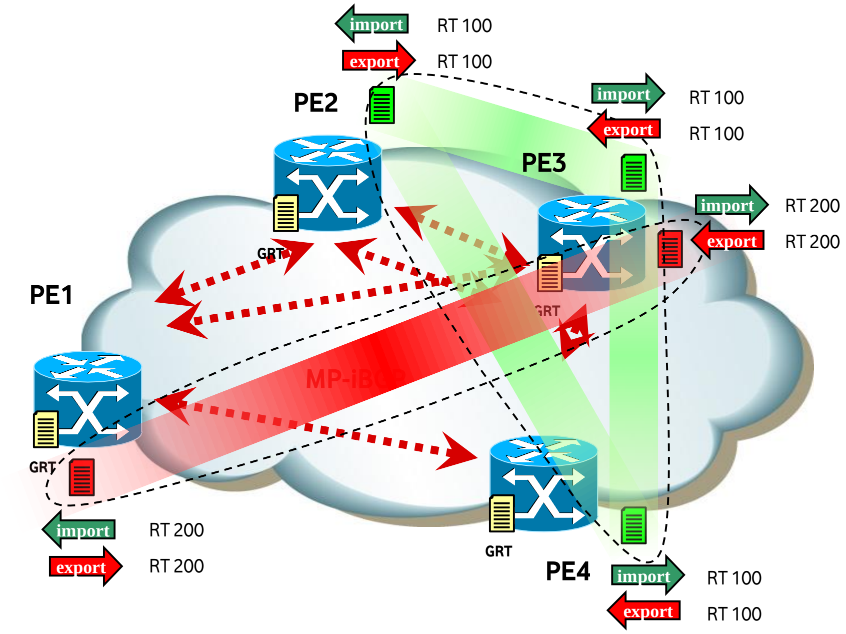
Each VRF has an associated (unique to the router) **Route Distinguisher (RD)**. The RD is used by MP-BGP to avoid confusing routes with overlapping IP addresses from different VPNs. It's not used to decide which route will be part of which VPN (the Route Target is used instead for this).
Even if it's common to use the same RD for "similar" VRFs on different PEs, this is not always the case.

**Each physical/logical router interface can be associated (at most) to one VRF**. By doing so that interface will be bound to the corresponding VRF. If a VRF is not specified, the interface will be bound to the Global RT.

One common situation is to have many CEs, each connected with a physical interface to the PE, with each interface associated to one of the PE's VRFs.

# VPN – VIRTUAL PRIVATE NETWORK

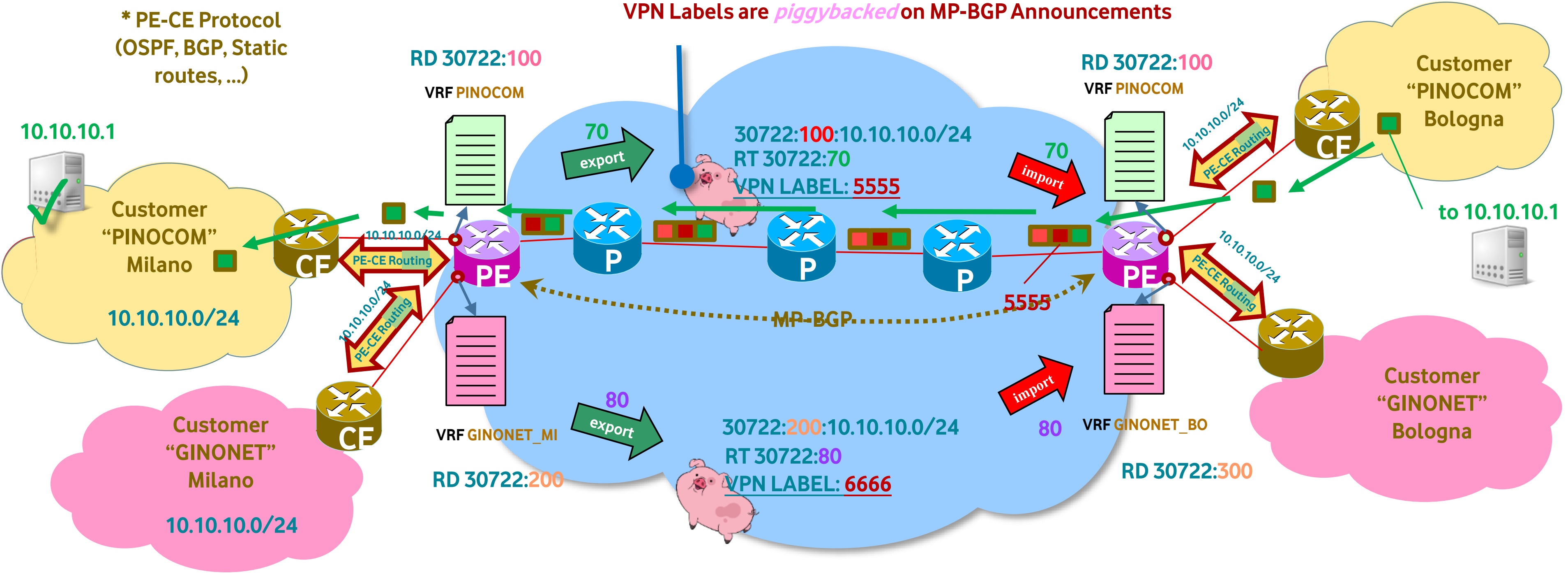
MPLS supports the creation of L3VPNs using Multi Protocol – BGP (MP-BGP) extension in a very flexible way



For example, using a RT value to denote a specific VPN, we can build full-mesh VPNs, completely isolated one from each other.

In this example we have **two full-meshed VPNs**, one associated with RT 100 and the other with RT 200.

## VPN Labels are *piggybacked* on MP-BGP Announcements



* PE-CE Protocol (OSPF, BGP, Static routes, ...)

RD 30722:100

VRF PINOCOM

RD 30722:200

VRF GINONET_MI

RD 30722:100

VRF PINOCOM

RD 30722:300

VRF GINONET_BO

10.10.10.1

Customer "PINOCOM" Milano

10.10.10.0/24

Customer "GINONET" Milano

10.10.10.0/24

Customer "PINOCOM" Bologna

to 10.10.10.1

Customer "GINONET" Bologna

70 export

80 export

70 import

80 import

30722:100:10.10.10.0/24
RT 30722:70
VPN LABEL: 5555

30722:200:10.10.10.0/24
RT 30722:80
VPN LABEL: 6666

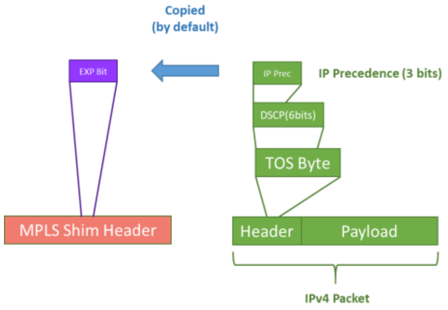
## Multi-Protocol BGP

Route Distinguisher
4 byte – To manage IP address overlapping
VPNv4 Address Family

Route Target – BGP Extended Community (4 byte) – To import/export routes in VRFs

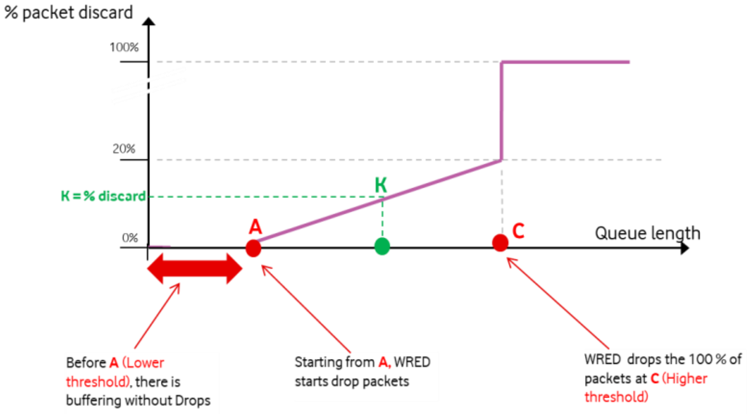
# QOS

- QOS class mapping in MPLS networks done using EXP bits.



| Class | DSCP | EXP bits | Queuing Algorithm | Scheduler |
|---|---|---|---|---|
| Control Plane | CS6, CS7 | 6,7 | - | PQ |
| Voice | EF | 5 | - | PQ |
| Enhanced/Standard | AF31, AF32, AF41, AF42 | 1,2,3,4 | WRED | PQ, CBWF, MDRR |
| Default | default | 0 | WRED | PQ, CBWF, MDRR |

- Three strict priority classes, for voice services and signaling, data traffic is assigned Default, Standard or Enhanced classes, with WRED algorithm ti handle congestion



Before A (Lower threshold), there is buffering without Drops

Starting from A, WRED starts drop packets

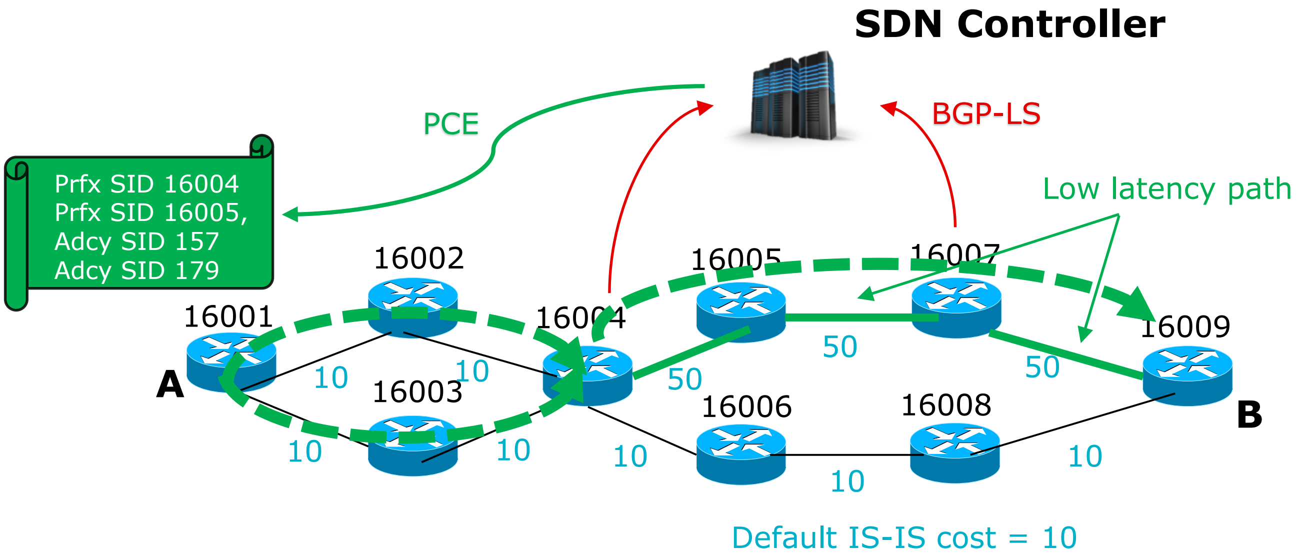
WRED drops the 100 % of packets at C (Higher threshold)

# SEGMENT ROUTING

- MPLS networks face a growing complexity in terms of variety of service requirements and scalability of LDP databases and number of tunnels.
- Segment Routing (SR) is a source-based routing: the source injects into the network the set of instructions to follow the routing path and encodes it in the packet header as an ordered list of segments
- SR can be directly applied to the MPLS architecture and integrates with multi-service capabilities including Layer 3 VPNs (L3VPN). A list of segments is encoded as a stack of MPLS labels.
- Segment IDs are distributed using IGP (IS-IS, OSPF) extensions only:
  - Prefix Ids, which uniquely identify a node (default SRGB 16000-23999)
  - Adjancency IDs which locally identify a link towards a neighbouring router
- No need of LDP or RSVP-TE to allocate Segment IDs or program forwarding information
- Traffic protection against link and node failures is faster (<50 msec convergence)
- Egress peering traffic engineering using BGP Segment IDs
- Dual  plane networks natively supported using Segment IDs anycast
- Plug&Play deployment thanks to interoperability with existing MPLS LDP dataplane

# SEGMENT ROUTING and SDN (Software Defined Network)

Segment Routing enables centralised traffic engineering,  agile programming source nodes only via Southbound Interface PCEP (Path Computational Element Protocol). No per flow state and signaling needed at midpoints and tail end routers



## Application Engineered Routing

- Segment IDs and topology info fed into SDN controller via BGP-LS
- Low latency service request from A to B
- Controller computes path and programs A with list of segments
- Equal Cost Multi Path using node prefix  SegmentID
- Low latency path selected using Adjacency SegmentID

# Agenda

1 Network overview: access network, legacy core network, NFV

2 IP backhauling, IP Core and underlying TX

3 IP MPLS and Segment Routing

4 SDN

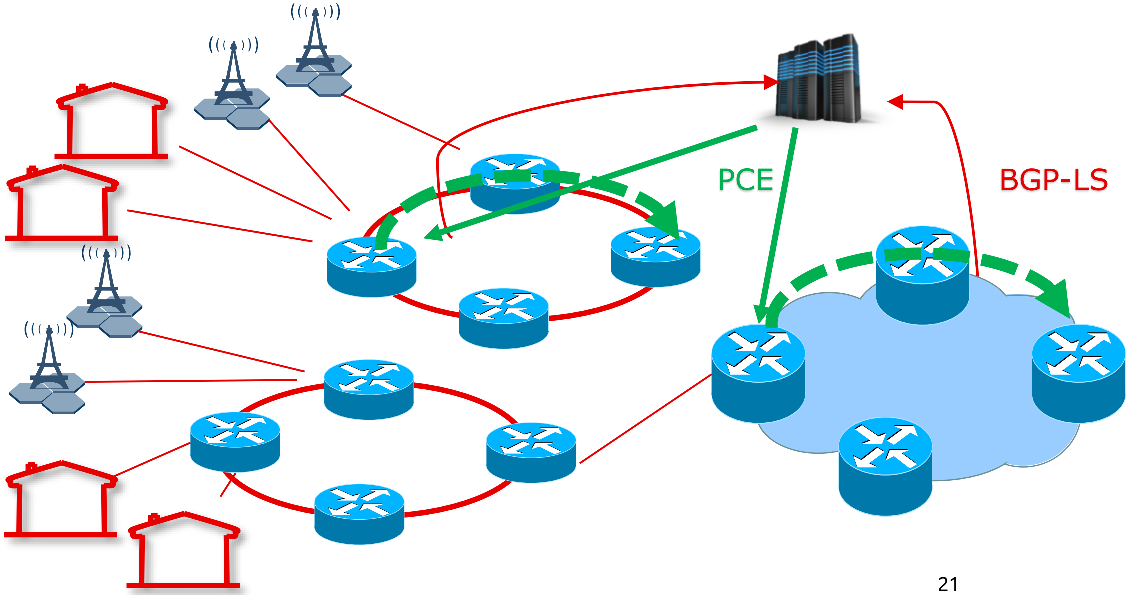
5 IP Network Operation Challenges & Evolution

6 Wrap-Up and Q&A

# Software Defined Network

SDN exposes transport network resources, supporting Network As A Platform architecture:

- Programmability, policy control, SLA fullfilment to support network slicing and service differentiation demand (capacity, latency, jitter, etc.)
- Automation, on line performance monitoring and planning tool, predictability to optimise network, improve resiliency and simplify operation



# Software Defined Network: use cases

Network and services autodiscovery:

Topology Discovery using BGP-LS

Dynamic network inventory

3° Party nodes control

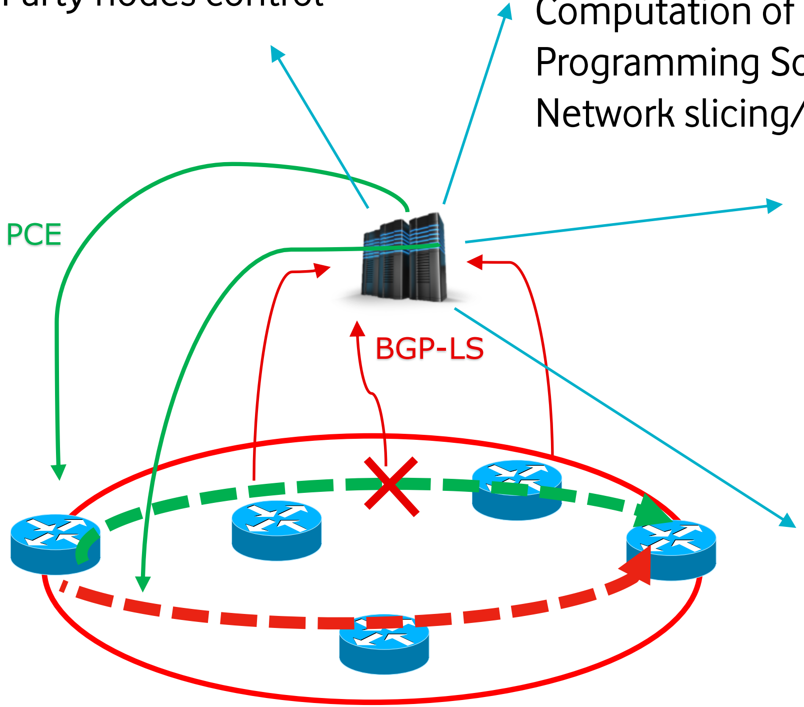
Service instantiation and provisioning

Service modeling using NETCONF/YOUNG

Computation of SLA adhering path and protection path

Programming Source nodes via PCEP

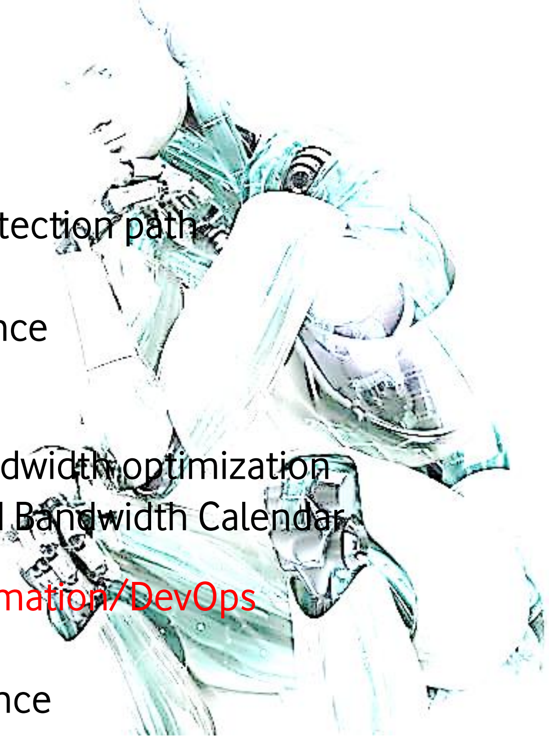
Network slicing/Disjoint Path/Path Avoidance



Network Optimisation

Capacity planning and bandwidth optimization

Bandwidth on demand and Bandwidth Calendar



Programmable automation/DevOps

Anomaly detection

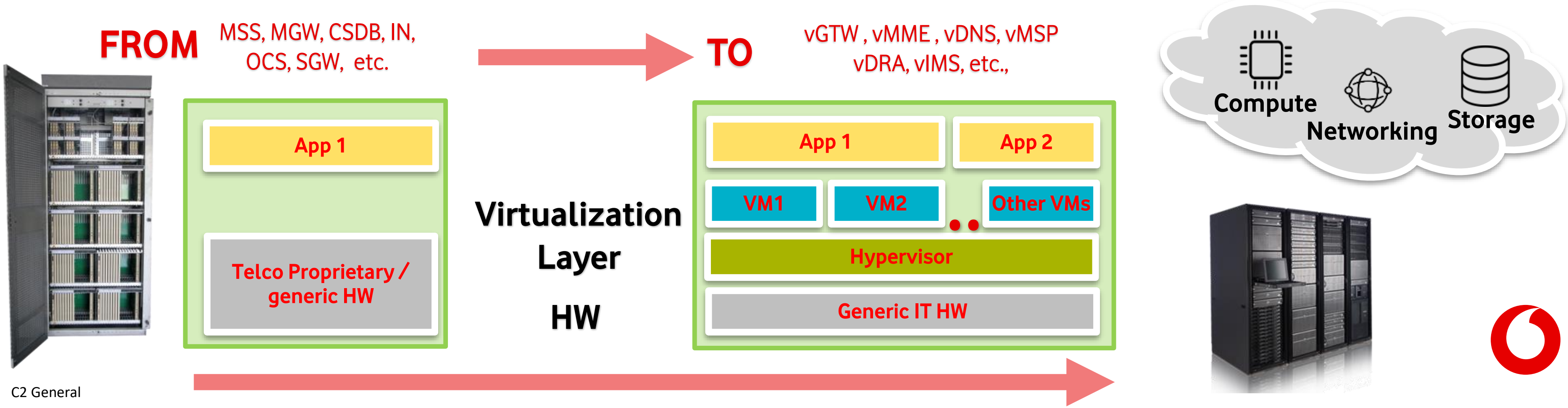
Predictive maintenance

What-if analysis

Dynamic congestion detection and alternative path creation

# SDN and NFV

- In the NFV architecture  netwok functions such as Evolved Packet Core, Switching, Firewalls, Baseband Processing Units, etc. are executed as software Virtual Network Functions instantiated in a shared pool of hardware, storage and computing resources, managed dynamically by a virtualisation software.
- Besides the advantages in terms of efficiency and flexibility to cope with a rapidly changing demand, NFV adds complexity to IP transport in managing the multiple traffic flows
- Moreover, 5G is a multi-service network, where ideally the physical network is «sliced» in isolated logical networks on a per service basis.
- SDN (combined with SR) provides a natural way to handle routing between VNFs through simplification and automation.



FROM

MSS, MGW, CSDB, IN, OCS, SGW,  etc.

TO

vGTW , vMME , vDNS, vMSP vDRA, vIMS, etc.,

App 1

Telco Proprietary / generic HW

Virtualization Layer

HW

App 1

App 2

VM1

VM2

Other VMs

Hypervisor

Generic IT HW

Compute

Networking

Storage

# Agenda

1 Network overview: access network, legacy core network, NFV

2 IP backhauling, IP Core and underlying TX

3 IP MPLS and Segment Routing

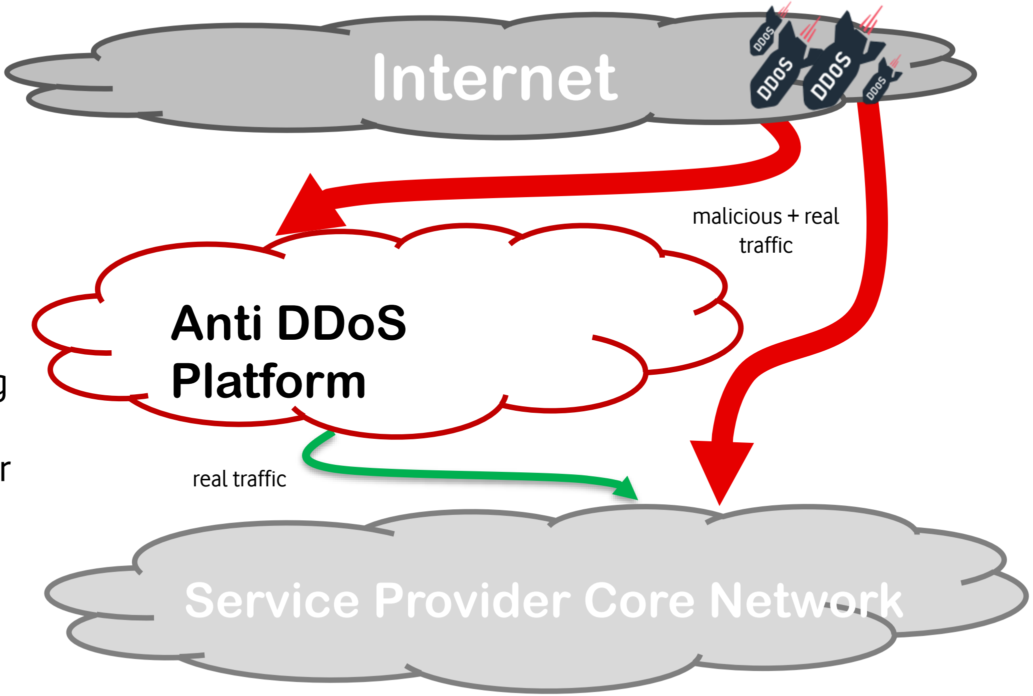
4 SDN

5 IP Network Operation Challenges & Evolution

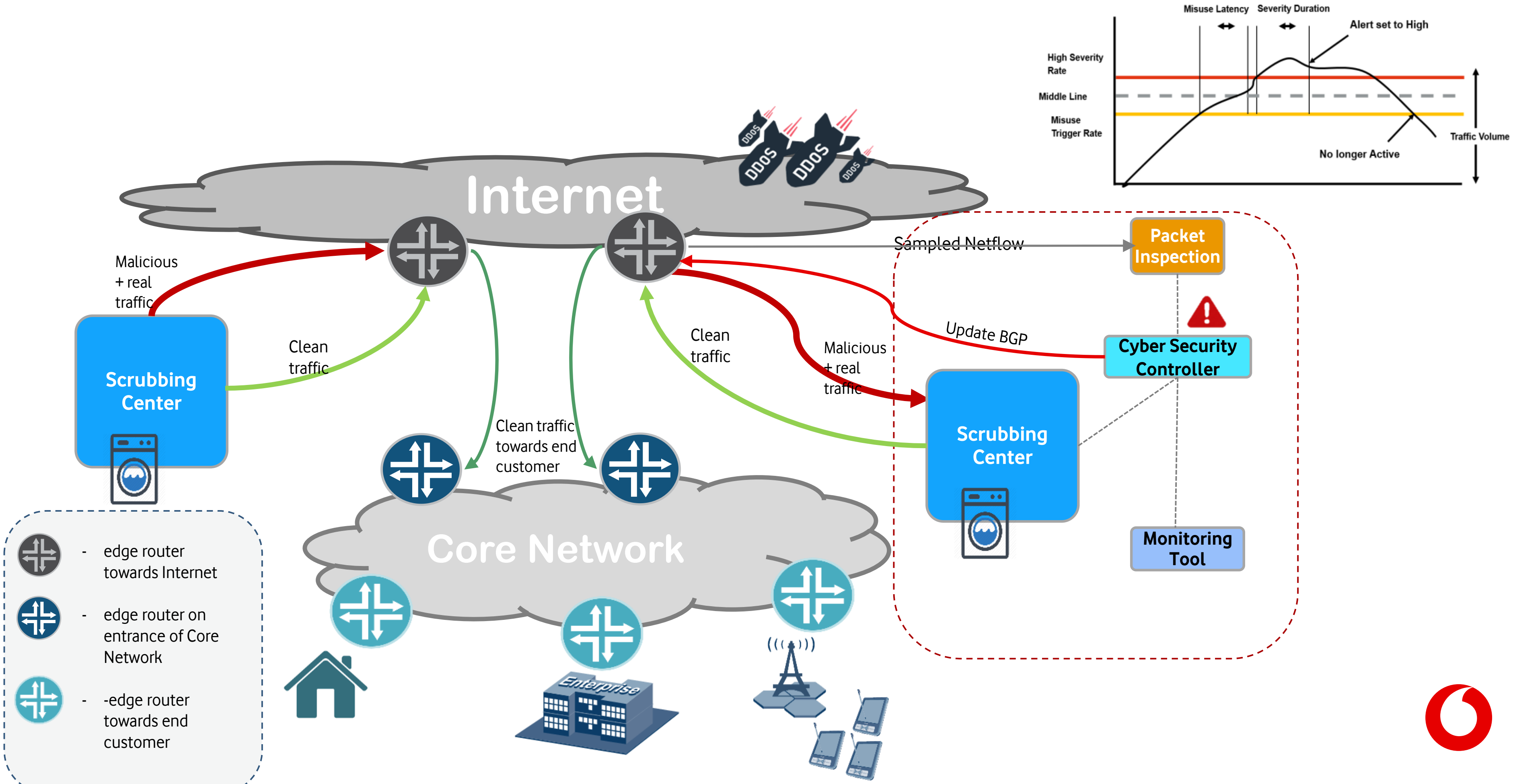
6 Wrap-Up and Q&A

# IP security aspects: An introduction to DDoS attacks

- DDoS (Distributed Denial of Service) malicious attempt to disrupt service using “Botnets”, networks of compromised “zombie” computers
- Types of DDoS attacks
  - Application layer attacks – ex. http flooding
  - Protocol attacks- ex. TCP SYN flood
  - Volumetric attacks – ex. DNS amplification
- Common DDoS attack strategies exploit different techniques to overload target servers  with flooding UDP, ICMP  (ping) traffic, or deviating traffic by announcing more specific routes (BGP hijacking), or causing disruption attacking NTP servers, or exploiting vulnerabilities still unknown or still without patching (Zero-Day Attack)
- Anti DDoS systems aim is to detect the attack, to scrub the malicious traffic



# Dedicated Anti-DDOS solution implemented



- edge router towards Internet

- edge router on entrance of Core Network

-edge router towards end customer

# Challenges: traffic trends and capacity management

- Data traffic growth is both technology and market driven
- Non linear effects play an increasing role in traffic profiles:
  - Gaming platform new releases
  - Simutaneous software upgrades downloads
  - Streaming of special events, football matches, concerts, etc.

⬇

- Cache servers increasing in number and moved towards the customer to relieve bandwidth requirements and costs on long distance
- Increase of direct peering points
- QOS!
- capabilities to support and enhance capacity planning processes

# Latency, latency, latency!

- Throughput is depending on latency!
- IoT applications rely on low latency
- Real time applications with very low latency requirements also to drive the future evolution of 5G mobile networks

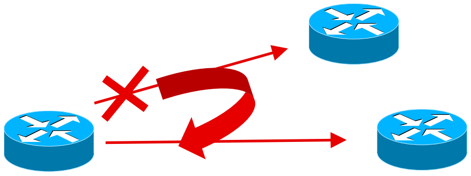
↓

- QOS is key to manage buffering and queuing privileging the low latency applications
- Application engineered routing
- Moving intelligence towards peripheral data center to shorten the physical path and reduce RTTs

# Resiliency (I)

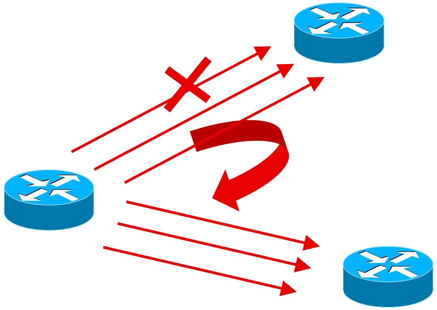
- Resiliency policy: active-standby versus load sharing

Active-standby optimizes latency, since primary and secondary path are likely to have different RTTs being on different physical paths. On the other end load sharing is in principle more efficient, and makes sure there is no unused link in the network minimising the risk of ‘silent’ issues



- Minimum-link configuration

Due to technical constraints primary/secondary paths may consist in 10/100Gbits link bundles. In case of a failure on a single link the primary path remains still active but with reduced capacity. The minimum-link parameter defines the minimum number of links active on the bundle which has to trigger switching on full capacity secondary path to avoid congestion



- Automatic re-configuration and intelligent re-routing to manage increasingly complex scenarios

# Resiliency (II)

- Critical components and risk analysis

Likelihood of double failures has to be carefully estimated, taking into account critical components failure rate and the expected time to recovery. For example islands connectivity via submarine cables shows a relatively low failure rate but a potentially high time to recovery

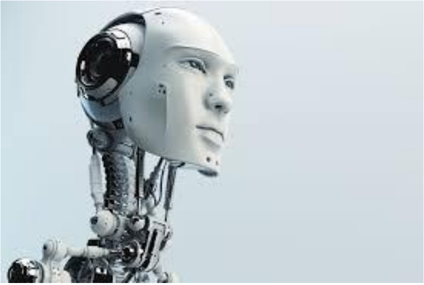
- Disaster recovery

Extraordinary events like heartquakes, flood, accidental fire may seriously impact service continuity of telco networks, which play strategic roles during an emergency situation. Specific countermeasures on site robustness, let alone recovery plans have to be designed and periodically tested to ensure Business Continuity

# Monitoring, preventive maintenance, robots (I)

- Real time monitoring is usually done looking at alarms propagated by network equipments. End to End service KPIs and traffic performance are also used though, being more effective to quickly identify the root-cause
- Extensive preventive checks on disturbances, event logs, misconfigurations do provide useful hints to prevent failures taking the proper actions proactively

- Automation is key to improve both reactivity and prevention, providing intelligent correlation engines, relieving humans in repetitive tasks and zeroing the risk of missing critical information



# Monitoring, preventive maintenance, robots (II)

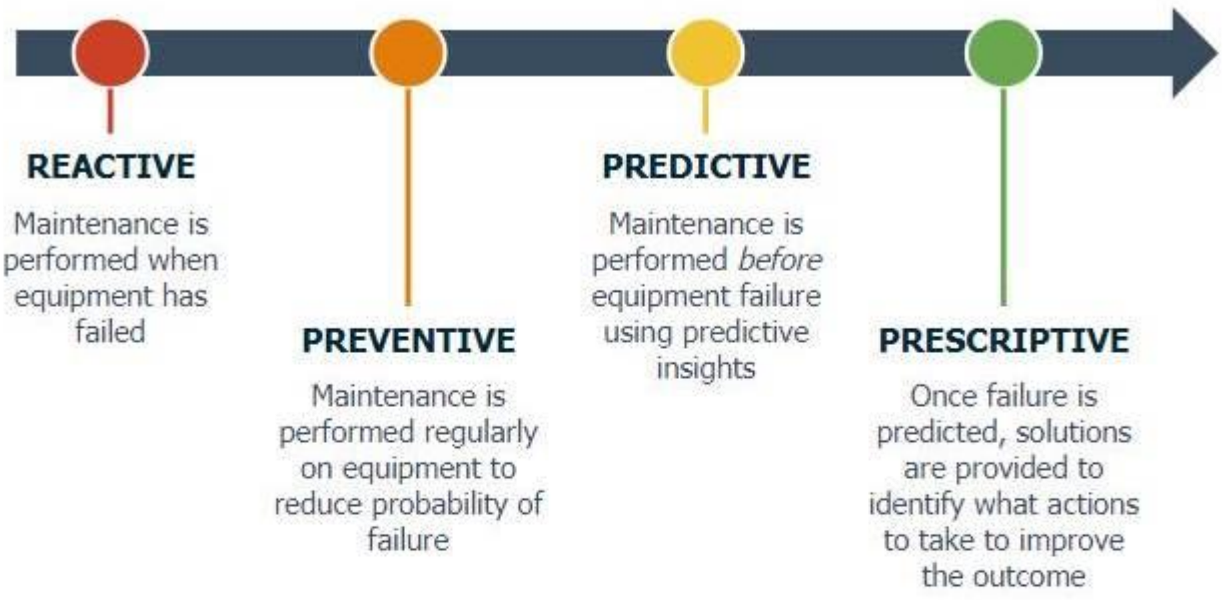
- Growing complexity makes increasingly difficult and risky manual tasks (i.e. manual correlation on VRF/VPN configuration data, or massive parameter changes in a large MPLS network )
- Traditional Network Inventory systems based on manual documentation could be hard to maintain and is prone to errors in matching data between network layers

↓

- Self discovery inventories, error free and automatically up to date, matching data between different layers
- The level of automation is of paramount importance to overcome the complexity in trouble-shooting, to minimize errors in configuration tasks

# From Preventive To Predictive

Extensive routine checks have been improving quality and reducing outage probability, Today exploiting big data and applying machine learning algorithms is expected to boost maintenance processes efficiency and achieve perfect quality.



**REACTIVE**

Maintenance is performed when equipment has failed

**PREVENTIVE**

Maintenance is performed regularly on equipment to reduce probability of failure

**PREDICTIVE**

Maintenance is performed *before* equipment failure using predictive insights

**PRESCRIPTIVE**

Once failure is predicted, solutions are provided to identify what actions to take to improve the outcome

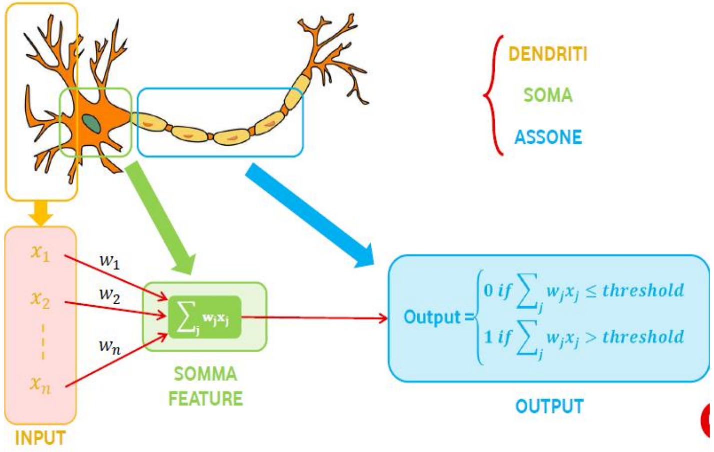
# Predictive Models

Deep learning/machine learning algorithms (Neural Networks, Random Forest, Vector Machine, Logistic Regression, etc.) are able to correlate events, identify patterns and make predictions

Precision, Recall, AUC KPIs used to estimate suitability to provide reliable predictions

Algorithms need training and fine-tuning.

Training could take advantage on expert hints rather than relying on a black box approach


DENDRITI
SOMA
ASSONE
$x_1$
$x_2$
$x_n$
$w_1$
$w_2$
$w_n$
$\sum_j w_j x_j$
INPUT
SOMMA FEATURE
$\text{Output} = \begin{cases} 0 \text{ if } \sum_j w_j x_j \leq \text{threshold} \\ 1 \text{ if } \sum_j w_j x_j > \text{threshold} \end{cases}$
OUTPUT

### PRECISION

$$\frac{\text{Nr. of events correctly predicted TRUE}}{\text{Nr. Of events predicted TRUE}}$$

$$\frac{5}{15} = 33\%$$

### RECALL

$$\frac{\text{Nr. of events correctly predicted TRUE}}{\text{Nr. of events actually occurred TRUE}}$$

$$\frac{5}{10} = 50\%$$

*Confusion Matrix*

| | | Prediction | |
|---|---|---|---|
| | | False | True |
| Reality | False | 80 | 10 |
| | True | 5 | 5 |

🟩 Reality and model coincide 🟥 Reality and model diverge

# From predictive to prescriptive

- Deep learning/Machine learning predictions are usually ‘blind’, they predict what will happen but cannot tell you why it will happen
- Algorithms can learn from humans and provide more insights on root causes
- Predictions should pave the way to proactive action: not an obvious step!
- Helping to govern an ever increasing complexity in an effective and efficient way is the ultimate goal

# In summary:

- IP networks are a key component of telco networks, they are growing in size and complexity, a growth that is going to pose considerable challenges in manageability, let alone security aspects.
- The flexibility of MPLS based networks, simplification with Segment Routing and programmability through SDN become an indispensable aid for both design and operation
- IP technologies do require from engineers high profile, very specialized skills. Nevertheless understanding the needs and the inclination to cooperate with experts from other areas, as well as the unrelenting thirst to learn new things will be the ultimate key for success

# Thank you!