TELECOM INFRA PROJECT®

Disaggregated Aggregation Router Technical Requirements

Authors

Kenji Kumaki, Ph.D.
Chief Architect & General Manager, KDDI Corporation
ke-kumaki@kddi.com

José Ángel Pérez
Principal Engineer, Vodafone
jose-angel.perez@vodafone.com

Ángel Cuadrado
IP Transport Engineer, Vodafone
angel.cuadrado@vodafone.com

Sarah Cook
Network Engineer, Orange
sarah.cook@orange.com

Nkosinathi Nzima
Senior Manager, Fixed Network Planning and IP Core, MTN
nkosinathi.nzima@mtn.com

Arturo Mayoral López de Lerma
Head of Transport Technology, Telecom Infra Project
amayoral@telecominfraproject.com

TIP Document License

By using and/or copying this document, or the TIP document from which this statement is linked, you (the licensee) agree that you have read, understood, and will comply with the following terms and conditions:
Permission to copy, display and distribute the contents of this document, or the TIP document from which this statement is linked, in any medium for any purpose and without fee or royalty is hereby granted under the copyrights of TIP and its Contributors, provided that you include the following on ALL copies of the document, or portions thereof, that you use:

1. A link or URL to the original TIP document.
2. The pre-existing copyright notice of the original author, or if it doesn't exist, a notice (hypertext is preferred, but a textual representation is permitted) of the form: "Copyright © 2021, TIP and its Contributors. All rights Reserved"
3. When space permits, inclusion of the full text of this License should be provided. We request that authorship attribution be provided in any software, documents, or other items or products that you create pursuant to the implementation of the contents of this document, or any portion thereof.

No right to create modifications or derivatives of TIP documents is granted pursuant to this License. except as follows: To facilitate implementation of software or specifications that may be the subject of this document, anyone may prepare and distribute derivative works and portions of this document in such implementations, in supporting materials accompanying the implementations, PROVIDED that all such materials include the copyright notice above and this License. HOWEVER, the publication of derivative works of this document for any other purpose is expressly prohibited.

For the avoidance of doubt, Software and Specifications, as those terms are defined in TIP's Organizational Documents (which may be accessed at https://telecominfraproject.com/organizational-documents/), and components thereof incorporated into the Document are licensed in accordance with the applicable Organizational Document(s).

Disclaimers

THIS DOCUMENT IS PROVIDED "AS IS," AND TIP MAKES NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT, OR TITLE; THAT THE CONTENTS OF THE DOCUMENT ARE SUITABLE FOR ANY PURPOSE; NOR THAT THE IMPLEMENTATION OF SUCH CONTENTS WILL NOT INFRINGE ANY THIRD PARTY PATENTS, COPYRIGHTS, TRADEMARKS OR OTHER RIGHTS.
TIP WILL NOT BE LIABLE FOR ANY DIRECT, INDIRECT, SPECIAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF ANY USE OF THE DOCUMENT OR THE PERFORMANCE OR IMPLEMENTATION OF THE CONTENTS THEREOF.

The name or trademarks of TIP may NOT be used in advertising or publicity pertaining to this document or its contents without specific, written prior permission. Title to copyright in this document will at all times remain with TIP and its Contributors.

This TIP Document License is based, with permission from the W3C, on the W3C Document License which may be found at https://www.w3.org/Consortium/Legal/2015/doc-license.html.

Table of Contents

Table of Contents (cont'd.)

01 Introduction

01 Introduction

This document represents the technical requirements for an open and disaggregated aggregation (DAR) device that operators can deploy in their backhaul transport networks. It describes the required hardware and proposes non-mutually exclusive software packages for the support of additional services or functionalities.

1.1 Why DAR?

The objective of this project is to develop a viable alternative to existing traditional packet transport solutions for deployment in Operator backhaul networks. From an operator perspective, many traditional solutions are:

- Monolithic in nature, which increases the complexity of introducing solutions or technologies from other suppliers
- Lack of openness in hardware to allow for different software stacks to run on top
- Lack of openness in software to allow for new or existing feature compatibility and extensibility
- Potential lock-in of device components such as pluggable modules
- Require vendor specific system integration

In this environment, operators find that current solutions:

- are based on monolithic platforms that make it extremely difficult to introduce innovation from other vendors in disparate parts of the device stack:
  - Lack of open hardware that can run various software types
  - Lack of open software that permits feature extensibility
  - Lack of fully open APIs that enable external components to interact with a device

This set of technical requirements aims to define an open and disaggregated platform that:

- is based on commercial, off-the-shelf components and open software that can perform routing and switch functions.
- provides the necessary scalability and resilience for various deployment scenarios.

1.2 Scope of the Document

The aim of this document is to describe:

- Target DAR platform architecture in relation to hardware and software feature set.
- Deployment scenarios where solution will outlining the relevant use cases and necessary features, functionality and capacity demanded by the operators.
- Hardware Requirements.
- Software Requirements.
- Operational management considerations.

02 Network Architecture and Solution Requirements

02 Network Architecture & Solution Requirements

This section describes the deployment models and typical network topologies where DAR will operate. Demanding on the scenario, elements such as port configuration/scalability and resiliency will have an impact on the hardware requirements of the solution and the associated software functionality

2.1 Traditional Transport Networks for Telecom Operators

There are many different topologies deployed and operated by different network operators. Our intention in this section is not to cover all this in detail. Instead, we start simply with the most common telecom operators network architecture, which can be illustrated with the example shown in Figure 1.

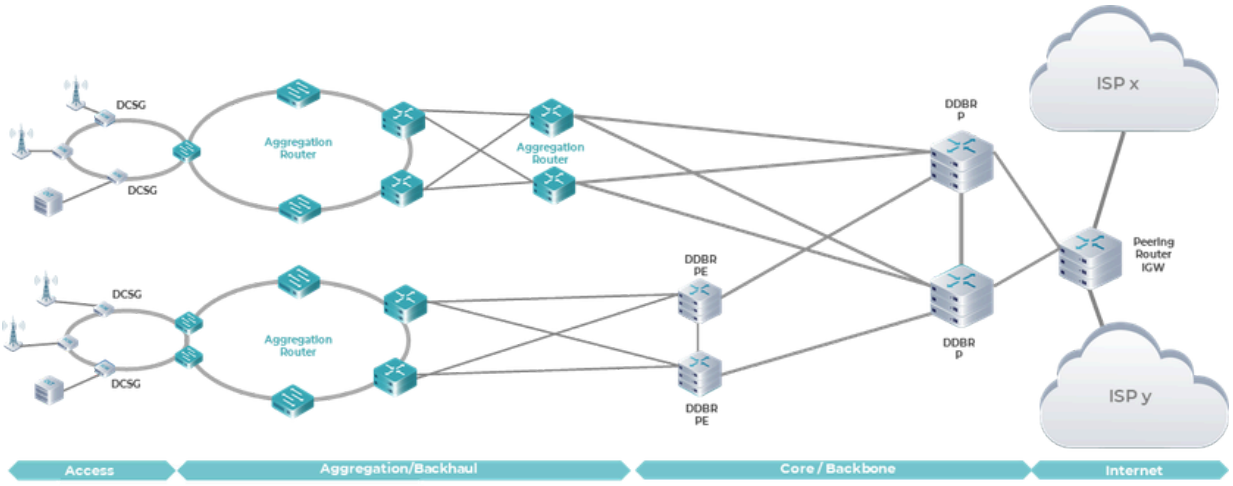


Figure 1. Traditional transport network segments for Telecom Operators

The network is subdivided into three segments: Access, Aggregation/Backhaul and Core. Network transport technologies utilized in each of these domains can vary from Layer 2 Ethernet, MPLS and Segment routing. In this case, DAR solutions may need to support multiple technologies at once, depending on the required services.

2.1.1 Access Network Aggregation

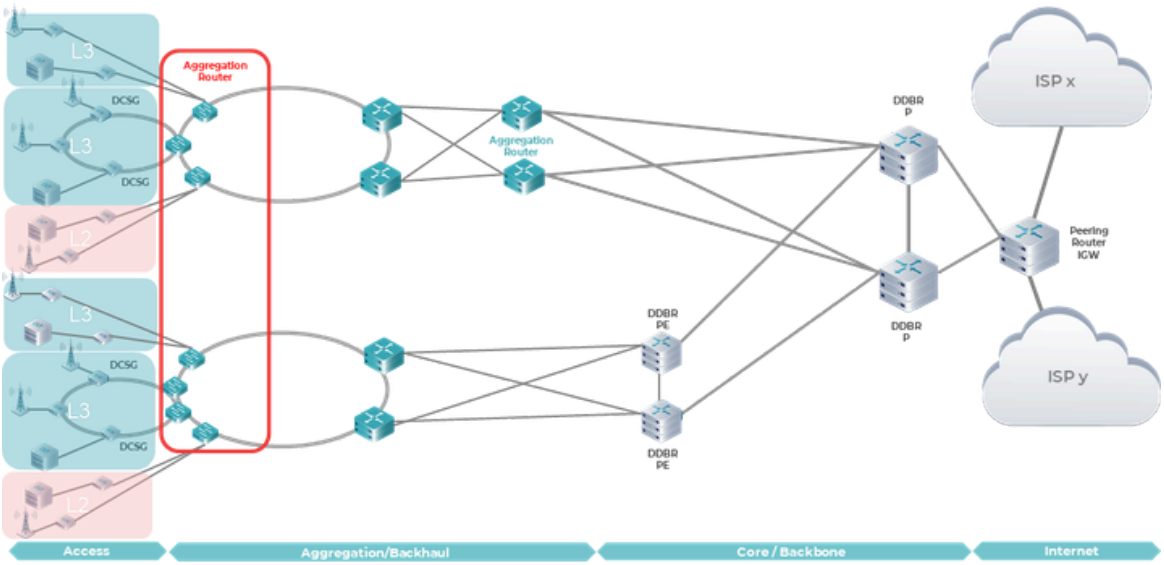


Figure 2. Access Network Aggregation

In this scenario, the DAR acts as the upstream device connecting one or more access networks up to the Core. This would include:

- Accommodating L2 and L3 Networks
- L3 VRF is used to accommodate multiple services such as Mobile, FTTH, and enterprise services
- Traffic Aggregation
- Traffic shaping for mobile traffic

2.1.2 Backhaul Aggregation

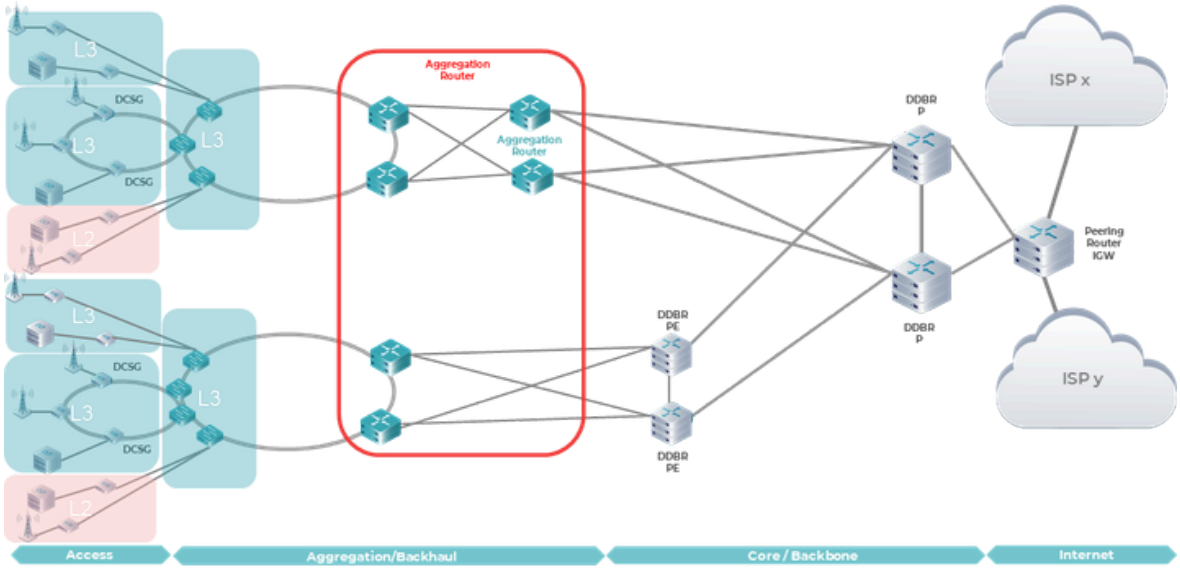


Figure 3. Backhaul Aggregation

In this scenario, the DAR interconnects additional downstream aggregation devices towards the access network domain up to the Core.

- Expected to only accommodate L3 networks.
- L3 VRF is used to accommodate multiple services such as Mobile, FTTH, and enterprise services.
- Traffic Aggregation.
- Traffic shaping for mobile traffic.

2.2 DAR Deployment Locations and Services

This section describes different deployment scenarios for DAR platforms and locations where they may be deployed. This describes specific scalability and functionality requirements as required.

2.2.1 DAR Deployment Locations

Operators may choose to deploy DAR in a variety of locations. The decision as to how widely DAR should be distributed depends on subscriber density, operational complexity and cost efficiencies.

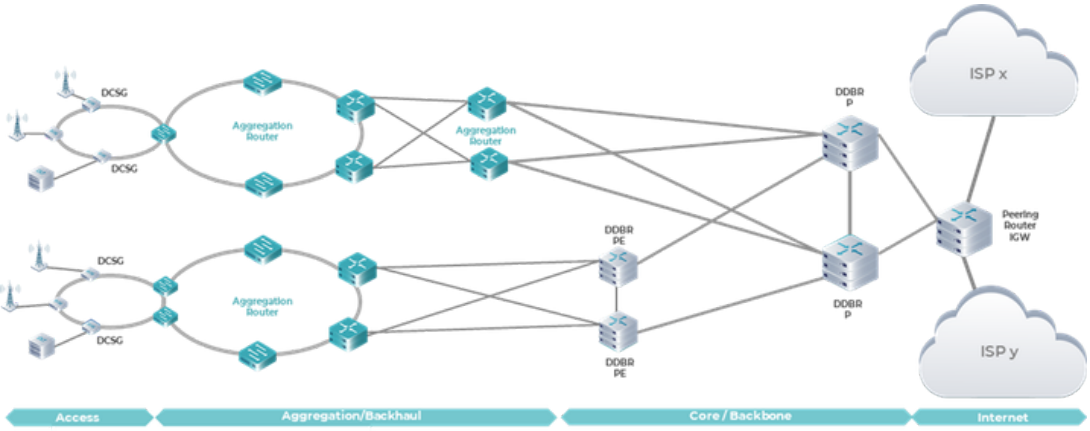


Figure 4. DAR deployment locations

Each deployment location has a different requirement for interface type, interface count and system resiliency. The characteristics of each location also mean that certain types of DAR solutions may be preferred over others.

2.2.2 DAR Deployment Models

DAR deployments will need to account for the following:

- Constrain the size of failure domains and associated 'blast radius' of any failures, to protect service levels and appetite for risk
- Have the ability to scale to a linear/uniform fashion
- Provide service resiliency to accommodate failure of one or more aspects of the solution; depending on traffic volume and service supported

This section describes the expected deployment models for DAR devices, which would meet the deployment scenarios already described.

2.2.2.1 Deployment Scenarios - Baseline

Typical deployments for DAR devices start with the baseline case:

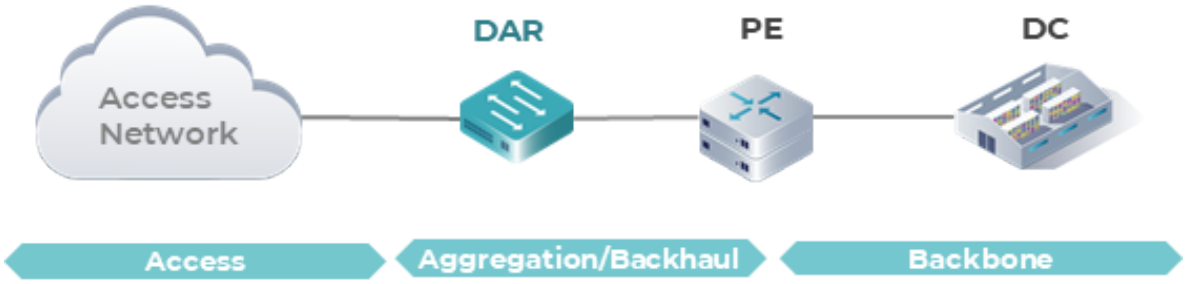


Figure 5. Baseline DAR Deployment

In this scenario, a DAR is deployed "in-line" between the downstream backhaul/access network and the upstream backhaul/backbone network. Though there may be multiple connections (utilizing link aggregation) between the DAR and wider network, in this baseline scenario the DAR could be a single-point-of-failure in the event that it fails.

2.2.2.2 Deployment Scenarios - Baseline (Redundant Node)

Building on the above, Telecom Operators could also choose a cluster or fabric-based node to deploy:

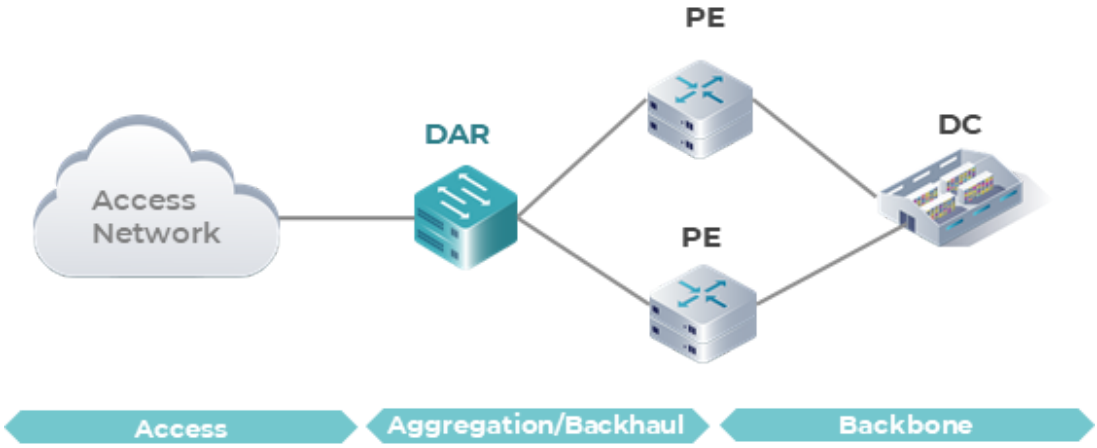


Figure 6. DAR Baseline Deployment - Cluster

Here, the DAR device is system resilient (e.g. consists of multiple routing/switching cards or similar), and as a result would be able to handle several first-order failure scenarios. Due to its design, the DAR node here may also be able to scale to higher capacities than offered by baseline solutions, though this is heavily dependent on ASIC use and resiliency design.

2.2.2.3 Deployment Scenarios – Horizontal Scaling

Building on the baseline case, capacity and service resiliency can be achieved with the following scenarios:

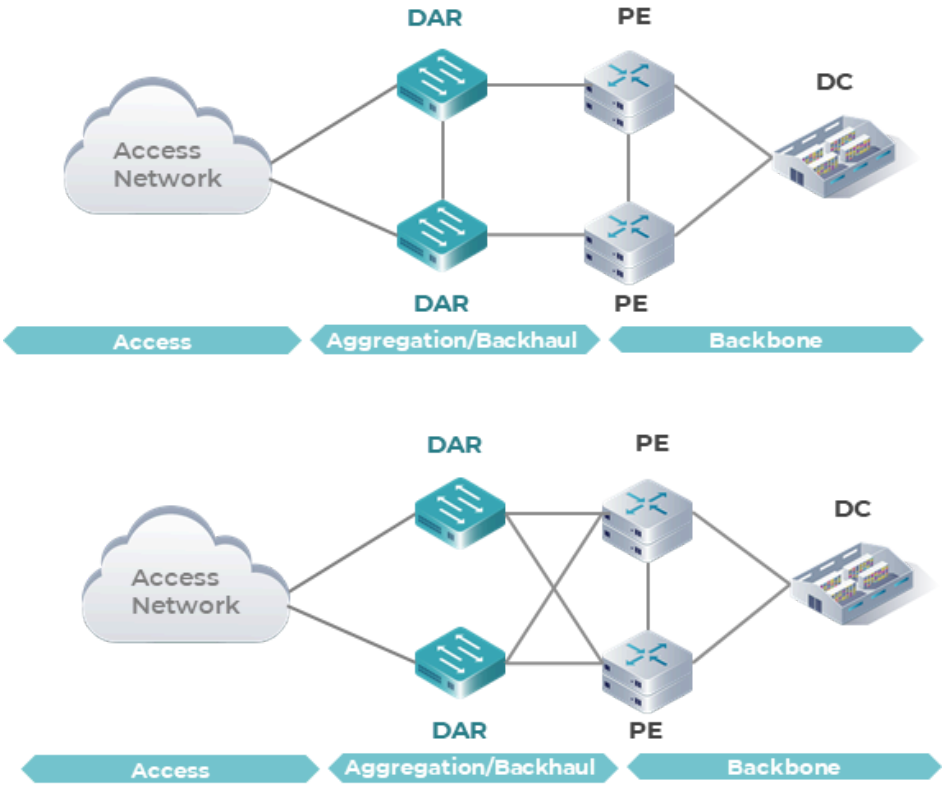


Figure 7. DAR Horizontal Scaling

In this case, DAR devices are deployed in a pair at the location, providing a layer of node redundancy. Depending on the requirements of the Telecom Operator, they have the flexibility to design the connectivity as they desire, with examples of East-West and multi-homing being depicted in the diagram above.

The solution could also be scaled to more than a pair of devices, though complexity will increase due to possible traffic balancing, segmentation and operating modes during failure scenarios.

2.2.3 DAR Services

This section is intended to describe the types of traffic flows and services that a DAR solution will be transporting across the network.

2.2.3.1 Mobile Access Aggregation

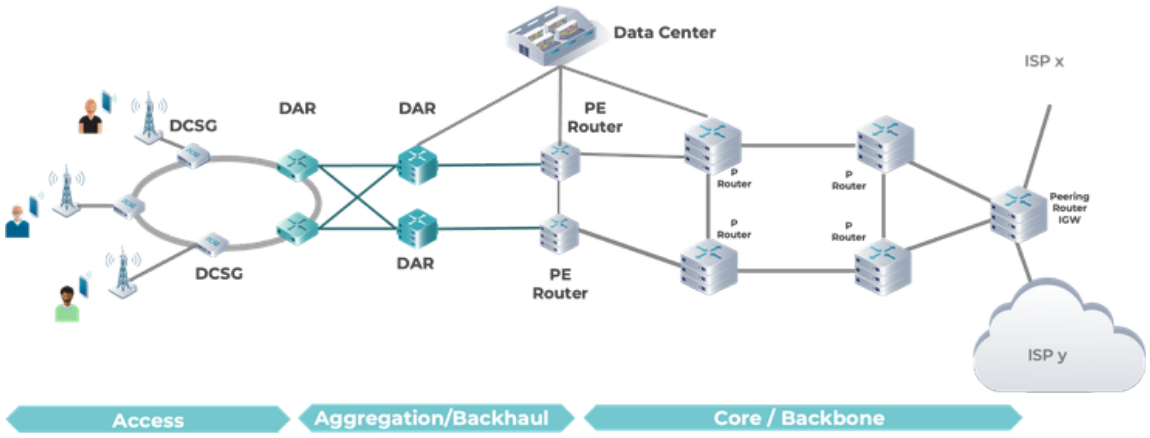


Figure 8. Mobile Traffic Aggregation

DAR devices will interconnect with DCSG devices in the access network to transport mobile traffic from cell site locations up to the core.

2.2.3.2 Cable Access Aggregation

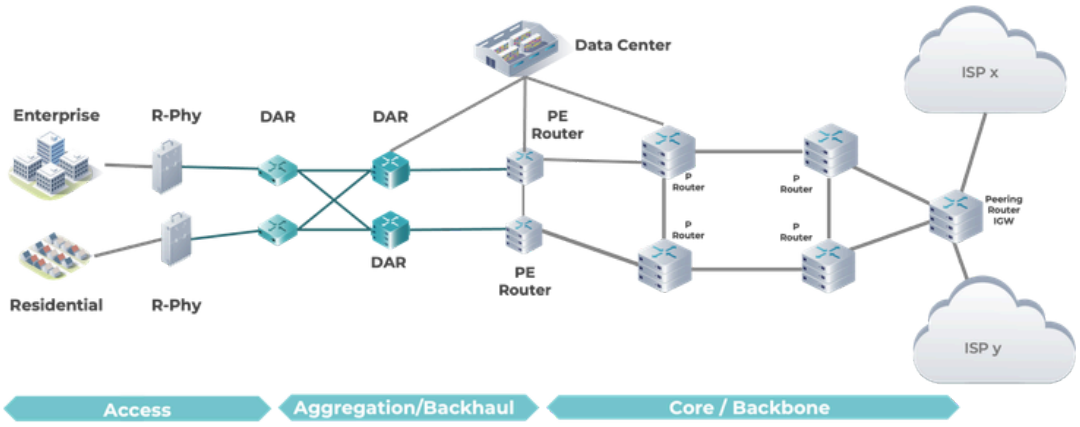


Figure 9. Cable Traffic Aggregation

DAR devices will interconnect with R-Phy cable devices in the access network to transport fixed line cable traffic from residential and enterprise locations up to the core.

2.2.3.3 PON Access Aggregation

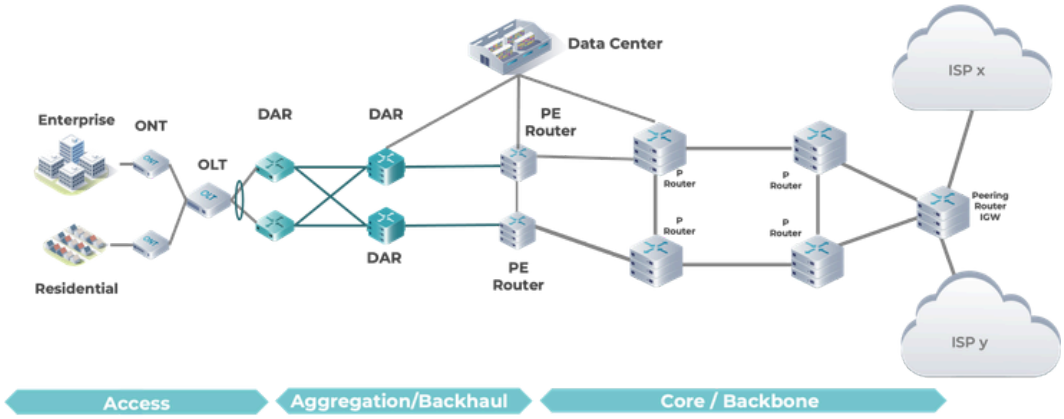


Figure 10. PON Traffic Aggregation

DAR devices will interconnect with OLT devices in the access network to transport fixed line cable traffic from residential and enterprise locations up to the core.

2.2.3.4 Fixed Wireless Aggregation

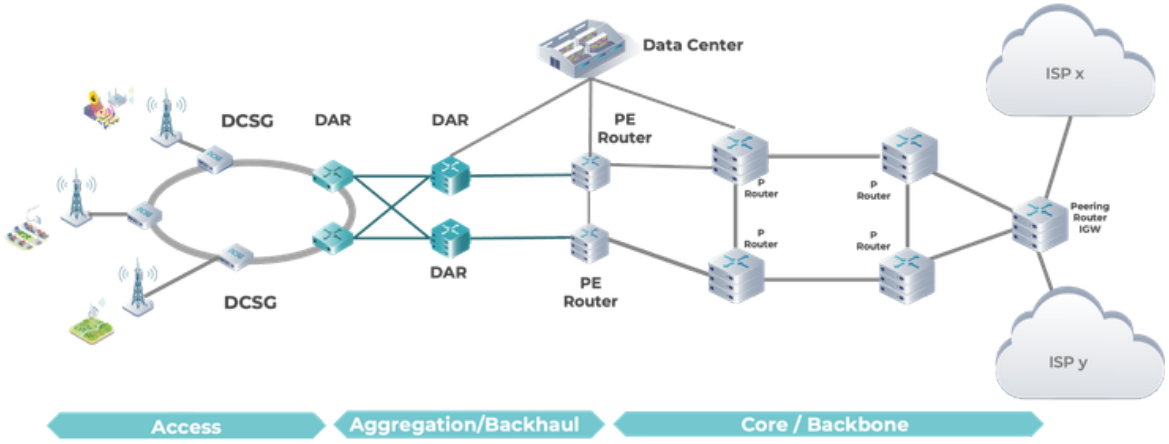


Figure 11. Fixed Wireless Traffic Aggregation

DAR devices will interconnect with DCSG devices in the access network to transport FWA traffic to the core.

2.2.3.5 Ethernet Access Aggregation

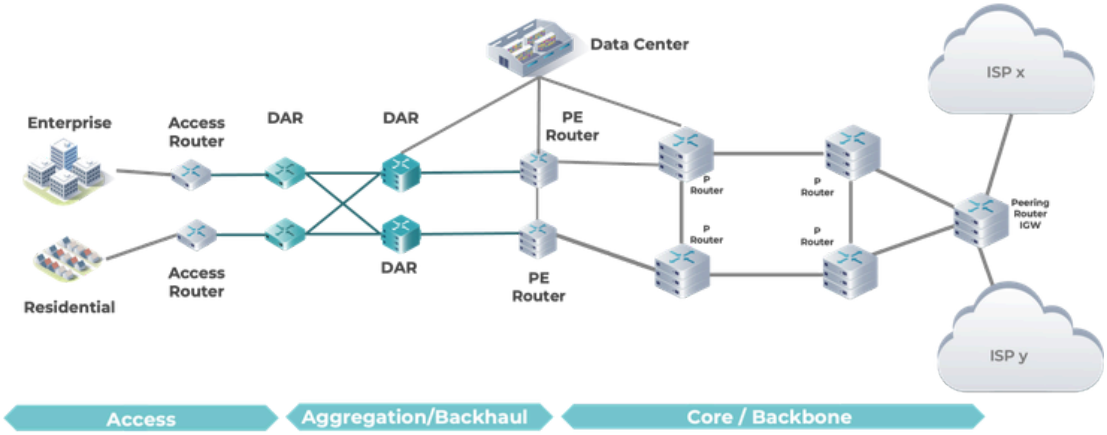


Figure 12. Ethernet Traffic Aggregation

DAR devices will interconnect with switches in the access network to transport fixed line traffic from residential and enterprise locations up to the core.

2.2.3.6 Combined Services Aggregation Summary

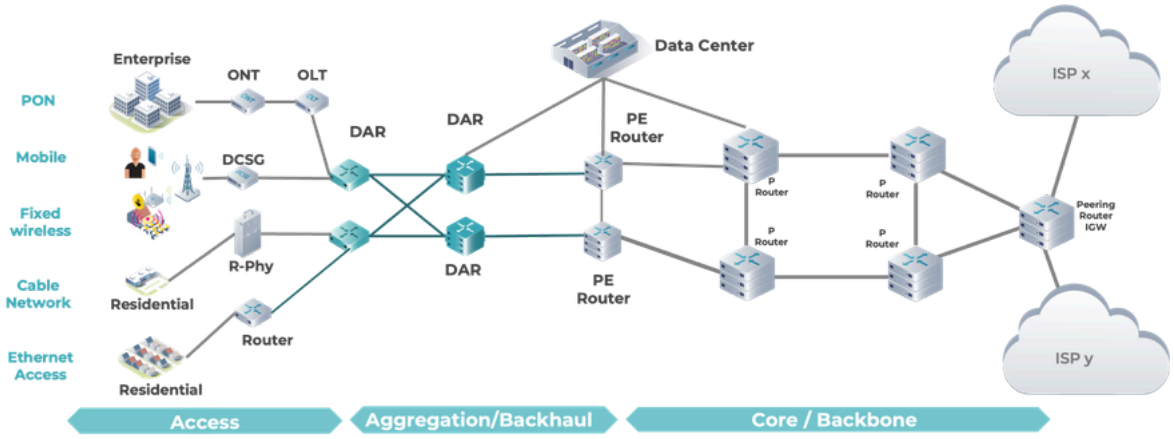


Figure 13. Combined Services Traffic Aggregation

To summarize, the DAR is expected to accommodate a combination of all these traffic types for transport from access to core. As such, it is important that the solutions provided are able to meet the necessary requirements of these services.

2.2.3.7 Datacenter Aggregation

Despite not being the core focus of this document, DAR solutions may also be considered for datacenter scenarios.

03 DAR Platform Architecture

03 DAR Platform Architecture

The main modules/components of the DAR platform are depicted in Figure 14 below.

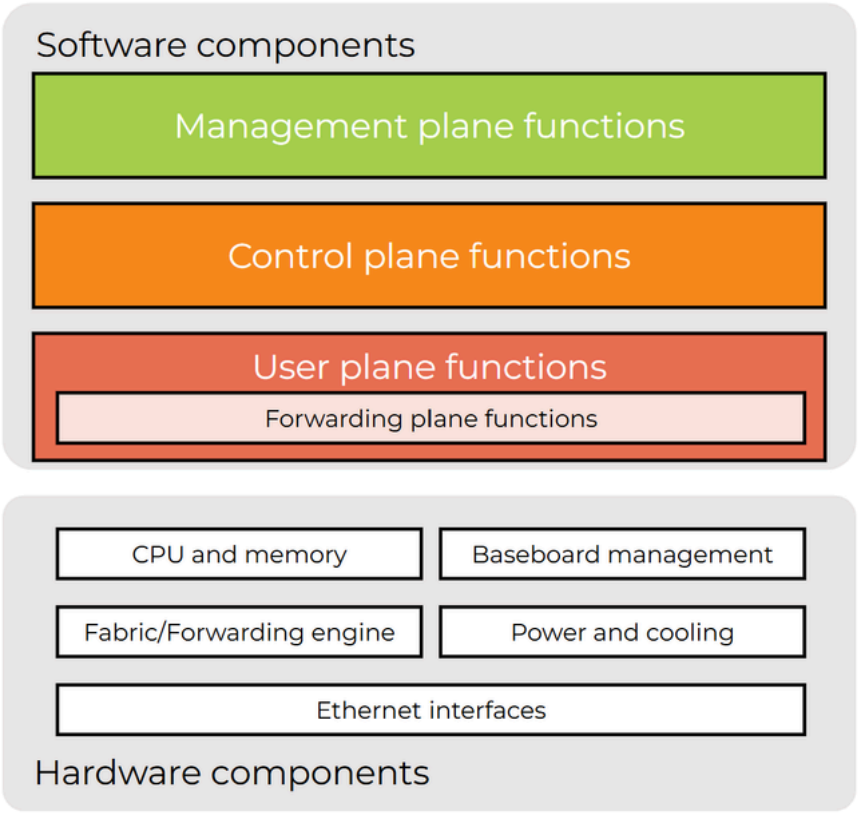


Figure 14. DAR platform architecture

The DAR node consists of:

- Commercial off-the-shelf (COTS) hardware with an open, programmable northbound interface (NBI). *Hardware requirements are described in section 4 below.*
- Open software (network operating system – NOS) with an open, programmable NBI. Software requirements are described in sections 5-7 in the current document.

Note that this technical requirements document does not define a strict position on how or where the software (or part of it) has to run (e.g. in the device, in an external VM). This is considered an implementation decision to be taken by vendors and network operators.

3.1 "Open and Disaggregated"

The 'open' in DAR does not imply 'open-source', but rather no 'hidden' or 'restricted' APIs that preferentially benefit the hardware vendor; or the NoS vendor; or the OS selected. DAR operators should be able to 'swap out' or replace any one of the three without affecting the other two.

To this end, the hardware system must not impose any restriction that limits the software that can run on it. In other words, the system must allow operators to install any operating system, even if its implementation comes from a third party. To ensure compatibility, it is highly recommended that Network Operating Systems for this platform are provided in the form of binary installers compatible with the Open Network Install Environment (ONIE) specification, as defined by the Open Compute Project (OCP). Equivalently, the DAR hardware will be equipped with ONIE.

If the platform provides the capability to verify the signature (via a particular certificate or a cryptographic key) of the software, it must be possible to disable such verification at any time, through software or firmware configuration, without the need for any specific or additional license. Moreover the hardware must be accompanied with a comprehensive toolkit of maintenance, configuration, diagnostic and repair information to facilitate modifications.

3.2 DAR Architecture

3.2.1 Standalone DAR

Standalone DAR solutions are based on single-chassis device solutions. In this case scalability and resiliency is achieved through the network topology design, in particular Spine & Leaf based architectures (Figure 15) are being considered by telcos to efficiently scale their DAR deployments and implement effective resilient topologies.

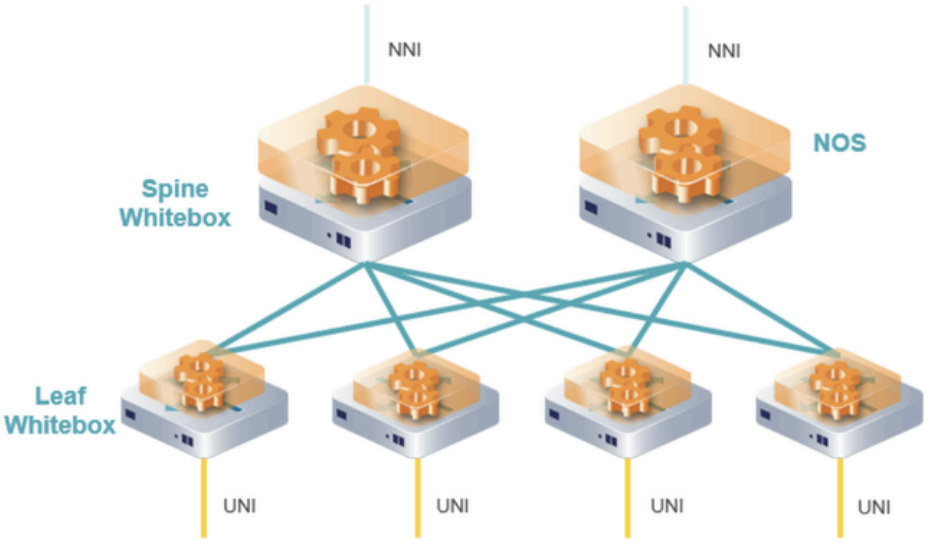


Figure 15. Disaggregated Spine and Leaf architecture.

Standalone deployments are intended for scenarios where traffic volumes and port requirements are limited in scope. It is expected that device capacity for standalone devices will range from 0.3Tbps to 10Tbps++. There may be scenarios where capacities greater than this range will be considered for standalone devices, however single-point-of-failure and resiliency concerns may become more prevalent.

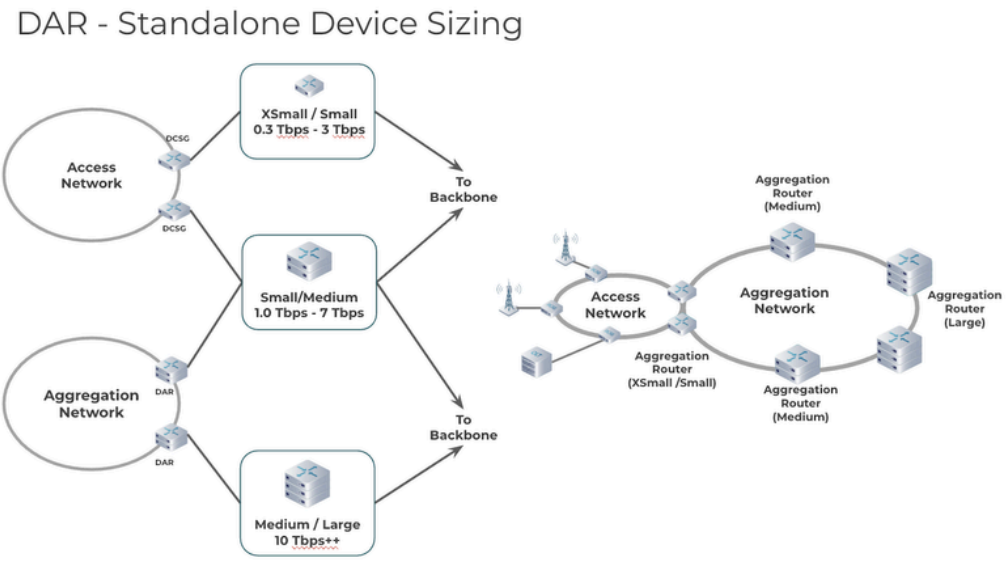


Figure 16. DAR standalone sizing

3.3 Scalability Figures

The table below provides a snapshot of the expected scalability figures of features for DAR solutions.
Please note that these are only reference figures, and more accurate ones will be provided in further detailed documentation:

Category	Topic	Large1	Large2	Medium1	Medium2	Small	XSmall
		12TBps+	12TBps+	7TBps -11TBps	3TBps -7Tbps	1.0TBps -3.0TBps	0.3TBps -1.0TBps
Interfaces	1 GBps Ethernet	0	0	0	0	0	8 to 32
	10 GBps Ethernet	0	0	40	40	40	8 to 24
	25 GBps Ethernet (optional)	0	0	0	10	4	2 to 24
	100 GBps Ethernet	40	40	10	40	4	2 to 4
	400 GBps Ethernet	20	20	20	6	1	0 to 2
Sub Interfaces	VLAN	200	1500	2000	4267	8000	8000
Bridge Domain	Bridge Domain	-	-	-	-	2000	2000
Neighbor Entry	Mac	100	1500	2000	128000	30000	30000
	ARP	100	1500	2000	42667	30000	30000
	IPv6 ND	100	1500	2000	85333	30000	30000
	DHCPv4/v6 (Interface)	-	-	-	4267	500	500
	DHCPv4/v6 (Client) ※SLAAC enabled for some base stations	-	-	-	21333	30000	30000
Shaping and Policing	Per System/Chassis	-	-	2000	85333	30000	30000
	Per 10GE PHY	-	-	1	160	1000	1000
	Per 100GE PHY	-	-	1	1560	10000	10000

Category	Topic	Large1	Large2	Medium1	Medium2	Small	XSmall
		12TBps+	12TBps+	7TBps -11TBps	3TBps -7Tbps	1.0TBps -3.0TBps	0.3TBps -1.0TBps
Number of Queues	Per System/Box	480	480	1280	256000	80000	80000
	Per 10GE PHY	8	8	8	480	250	250
	Per 100GE PHY	8	8	8	4680	2500	2500
Number of ACLs	Per System/Chassis	-	200	10000	85333	8000	8000
	Per PHY	-	50	1000	1560	3000	3000
Number of Neighbours	OSPFv2	50	50	10	50	-	-
	OSPFv3	50	50	10	50	-	-
	IS-IS	-	-	-	0	20	20
	MP-BGP	400	1000	400	100	50	50
	Multicast (IPv4)	50	50	-	0	20	20
	Multicast (IPv6)	-	-	-	0	-	-
	LDP (connected)	300	-	10	0	-	-
	LDP (targeted)	200	-	-	0	-	-
	RSVP	300	-	-	0	-	-
	SR	300	-	-		20	20
	VRRP ※Large has multi-VRRP (Two groups are running in one VLAN)	-	-	2000	4267	-	-
	BFD	200	200	30	4267	500	500
	PIM	-	100	-	-	30	30

Category	Topic	Large1	Large2	Medium1	Medium2	Small	XSmall
		12TBps+	12TBps+	7TBps -11TBps	3TBps -7Tbps	1.0TBps -3.0TBps	0.3TBps -1.0TBps
Number of Prefixes	OSPFv2	11000	0	11000	200	0	0
	OSPFv3	6000	0	6000	2000	0	0
	IS-IS	-	0	-	-	1000	1000
	BGP (IPv4)	2M	200K	2M	2000	30000	30000
	BGP (IPv6)	1M	100K	1M	2000	30000	30000
	Multicast (IPv4)	-	-	-	-	-	-
	Multicast (IPv6)	-	-	-	-	-	-
	Static (IPv4)	1000	500	1000	2000	1000	1000
	Static (IPv6)	1000	500	1000	2000	1000	1000
	Subscriber (DHCPv6 PD)	-	-	-	-	30000	30000
OSPF	Area	-	-	-	20	-	-
	Multi-Process	-	-	-	5	-	-
OSPFv3	Area	-	-	-	20	-	-
	Multi-Process	-	-	-	5	-	-
Numbers of Groups	Multicast (IPv4)	1000	500	-	-	100	100
	Multicast (IPv6)	-	-	-	-	-	-

The total capacity
** : In case breakout cables are deployed, the maximum number of 10GE is 160.*

Table 1. DAR Scalability

Category	Topic	Large1	Large2	Medium1	Medium2	Small	XSmall
		12TBps+	12TBps+	7TBps -11TBps	3TBps -7Tbps	1.0TBps -3.0TBps	0.3TBps -1.0TBps
VRF	VRF	100	100	-	10	10	10
MPLS-TE Path	RSVP Path (Head/Mid/Tail)	300/600/ 300	-	-	-	-	-
	VRF	100	-	-	-	-	-
Load Balancing	LAG	32	20	32	32	32	32
	ECMP	32	20	32	32	32	32

Table 1. DAR Scalability

04 Hardware Requirements

04 Hardware Requirements

The following sections describe the requirements of DAR hardware platforms, including environmental characteristics, performance expectations and timing/synchronization requirements.

4.1 Hardware Solution Form Factor, Power, Cooling and Environmental Conditions

The DAR hardware should consist of the following:

- 1-3 RU form factor, suitable for deployment in telecoms standard 19-inch rack
- Front-to-back airflow, though others such as side-to-side may also be considered
- AC and DC power supplies operating in a 1+1 redundancy fashion with the ability to hot swap
- Hot swappable fan modules operating in a redundant configuration
- Operating temperature ranges from 0-45 degree Celsius

4.2 Hardware Platform CPU and Forwarding Engine

It is expected that DAR platforms would comprise ASICs which are designed for high-performance networking applications specifically for network edge and aggregation domains. Interoperability factors should also be taken into account when specifying solutions

4.3 Hardware Platform Management

The DAR hardware must include, as a minimum, one console and one management port (both RJ45) and one USB 3.0 port, for local configuration and debugging. Nonetheless, it must be possible to remotely disable the console and/or the management port and/or the USB port. It is therefore mandatory that the platform also supports in-band management.

The hardware platform must include status indicators, including per port LEDs.

4.4 Hardware Platform Management

As stated in section 2, the DAR hardware may be used to aggregate MBH. In these cases, the DAR shall be able to propagate synchronization signals to other network elements directly or indirectly connected to it. Furthermore, it may also need to provide frequency, time and phase synchronization to 2G/3G/4G/5G base stations directly connected.

The following are the mandatory synchronization requirements:

- Support of IEEE 1588 profile as defined in ITU-T G.8275.1 (full-timing support; Telecom - Boundary Clock) and Sync-E for holdover purposes and Grandmaster redundant sources support. The requirement corresponds to the support of IEEE1588v2 – Precision Time Protocol – profile for telecoms (multicast mode preferred) and includes Sync-E in Ethernet interfaces as per ITU-T G.8261 (section 9.2.1), G.8262 and G.8264.
- Network quality model (microsecond precision) according to ITU-T G.8271.1.
- Node performance (noise generation, tolerance, transfer and holdover) according to ITU-T G.8273.2 (sections 7.1/7.2/7.3/7.4).
- Node performance (upon wander, failure and holdover) according to ITU-T G.8273.2 (sections 7.2/7.3/annex). Clock type class B (minimum) or C (desirable), as defined in ITU-T G.8273.2.
- Support of IEEE 1588 profile as defined in ITU-T G.8275.2 (partial-timing support).
- Network quality model (microsecond precision) according to ITU-T G.8271.2.
- Node performance (noise generation, tolerance, transfer, wander, failure and holdover) according to ITU-T G.8273.4 (sections 7 & 8). Clock type class B, as defined in ITU-T G.8273.4.

The electronics used to build the synchronization regeneration capabilities are to be selected by the platform manufacturer (ensuring full performance compliance with standards), but detailed information should be shared within TIP. Hardware SKUs should include at least 2-time synchronization solutions to allow software (SW) providers to implement the above-mentioned features using their preferred SW stack.
Finally, the hardware platform must include, at least, one GPS signal input interface (which may be based on SFP) and one 1PPS interface for external synchronization. These will be used in scenarios where the synchronization signals cannot be received from the network, and there are base stations directly connected. It may also be possible to select one signal among several (by means of a BMCA, desirably hardware implemented), in case there are multiple sources.

4.5 Hardware SKU Network Interfaces and Forwarding Capacity

The required capacity and the interfaces layout in the hardware platform heavily depend on the specific DAR scenario (i.e. platform type, deployment location and deployment model) that is being considered. General requirements will be included first, with specifics per scenario defined later.

The solution must be able to support electrical and optical interfaces as per IEEE 802.3, and they shall be configurable to work either as UNI or as NNI. Pluggable optics will be preferred to fixed format connectors: they shall be able to operate at the same temperature ranges as the node, and it will be possible to configure them at different speeds (e.g. 1G/10G with SFP+) without reboot. There will be no limitations on the type of connectors that are used (SR, LR, etc.; LAN/WAN PHY), and the platform must be fully interoperable with third party optics. Additionally, the system shall be compatible with third-party coloured WDM pluggable optics (tuneable & fixed).

All physical network interfaces in the proposed platform must support multiple services simultaneously, independently on whether they are: multiple PWs each with its own set of VLANs; multiple VLANs associated to VRFs, VPLS, E-VPN, etc.; multiple native VLANs or VLANs associated with core interfaces (those configured with IS-IS, OSPF, LDP, etc). Similarly, ethernet LAG interfaces must support the same type of services as single physical interfaces.

4.5.1 Interface Types and Requirements

It is expected that the following interface types will be required for DAR solutions:

- 1000BASE-SX
- 1000BASE-LX
- SFP-10G-LR
- 100G-LR4
- 100G-ER4
- 400G-LR4
- 400G-ER4
- 400G-ZR/ZR+

The 1G/10G/25G Ethernet port cages should support both short-reach (SR) or long-reach (LR, ZR, ER) pluggable optical transceivers, and direct attach copper (DAC) transceiver-cable assemblies. In addition to downstream connections to access nodes, it is expected that a number of these ports will be used (1G/10G/25G) for local connections to servers hosting control and management plane as well as other services like wholesale interfaces or legal intercept connections. The use of breakout cable to provide these ports is discouraged as these would limit the flexibility to support different configurations of optical transceivers.

For 100G/400G ports, again there is a mixture of interfaces for short-reach (SR) and longer distances (LR, ZR, ER). These interfaces provide connections between the leaf and spine switch (100G) as well as connections to the network core (100G/400G).

With regards to the forwarding capacity, the main requirement is to present non-blocking forwarding architectures, as already commented in section 3.2. Note: Oversubscription of the spine is expected, as leaf nodes will be run with less capacity for redundancy reasons.

05 Software Requirements

05 Software Requirements

The following list of requirements is included to ensure that the DAR platform will be able to work in any type of network scenario potentially deployed in the production networks of the participating operators. Some of these requirements may be considered as prerequisites for the support of VPN services (i.e., pre-requisites for the DAR operating as a PE). Examples of these may be MP-BGP or T-LDP. However, they will be included here to avoid too much dispersion in the location of requirements.

5.1 Layer 2 Switching

A high level list of layer 2 switching requirements include:

- Layer 2 forwarding and bridging
- VLANs (802.1q, 802.1ad etc)
- Spanning Tree Protocol
- Bridge Domains
- LACP (described in more detail in this chapter)
- Jumbo frames

5.2 IP/MPLS Routing

Firstly, it is important to note that the DAR must support dual stack (IPv4 and IPv6), or IPv6 only, in all its network interfaces[1]. In fact, for all the protocols mentioned below, it will be understood that they must support the applicable IPv6 extensions (when applicable). The high-level list of routing protocols is the following:

- Static routing.
- IS-IS, including extensions for Traffic Engineering.
- OSPFv2/v3, including extensions for Traffic Engineering.
- BGP-4, including multiprotocol extensions, capabilities advertisement (RFC5492), communities (RFC1997), BGP-LU (RFC3107), deterministic-med (RFC 4721), graceful restart / non-stop Forwarding (RFC 4724), extensions for 4-byte AS number (RFC4893), confederations (RFC3065), route reflection (RFC4456), error-handling (RFC 7606), peer tracking (RFC 7854) and prefix-Independent Convergence (PIC)[2].

[1]Requirements for interfaces towards the CE, when the DAR acts as a PE node, are defined in section 7.

[2]Bashandy, A., Ed., Filsfils, C., and Mohapatra, P.: "BGP Prefix Independent Convergence", Work in Progress, https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-bgp-pic-13, February 2021.

With regards to these protocols, the proposed platform must support distributing routes between any of them, based on defined policies, and also distributing local and static routes. It must also support modifying the priority (administrative distance) of the different routing protocols when populating the active FIB. Any mechanism for mutual authentication must support at least MD5 authentication.

MPLS must be supported as well, with at least four MPLS labels in the label stack. Signaling of MPLS labels must be possible both using:

- LDP, including T-LDP and IGP and LDP synchronization
- RSVP-TE, with the capability of path computation based on CSPF and supporting FRR mechanisms
- Segment Routing

Finally, other features that are required include:

- ECMP
- Entropy-label-based load balancing[3]
- IP LFA FRR mechanism
- VRRP
- Multicast capabilities, both for IPv4 and IPv6
- PWE3

Segment Routing (and other mechanisms like TI-LFA) shall be supported, using MPLS encapsulation (SR-MPLS) and associated OSPF/IS-IS extensions. Support for alternative segment routing technologies (e.g. SRv6) would be valued. If segment routing is not supported, then proposals complying with this specification must demonstrate a clear roadmap to implement Segment Routing. Policy-based routing will be required to eliminate forwarding to non-adjacent next-hops i.e. a mechanism to prevent subscriber-to-subscriber traffic flow. On the other hand, the DAR must support PCEP and BGP-LS, to account for the deployment of PCE nodes together with RSVP-TE.

Finally, and despite the strong focus on layer-3 functionalities, support of layer-2 Ethernet switching is also required in the DAR.

[3]Akiya, N., Swallow, G., Pignataro, C., Malis, A., and Aldrin, S.: Label Switched Path (LSP) and Pseudowire (PW) Ping/Trace over MPLS Networks Using Entropy Labels (ELs); https://datatracker.ietf.org/doc/html/rfc8012, November 2016.

5.3 Link Aggregation

The proposed platform must support aggregating several physical ports into a single logical interface, based on LACP protocol as defined in IEEE 802.1ax. Both LACP active-active and active-standby shall be supported.

Schemes for balancing the load among the different ports that achieve a balanced share will be preferred. It shall also be possible to determine by configuration the number of interfaces within a LAG that determine the failure status of the whole logical interface (e.g. 1 out of 3, 2 out of 2...).

All the IP/MPLS, QoS, synchronization, OAM and Performance Monitoring functionalities shall work on logically aggregated interfaces.

The traffic balancing algorithm must be based on MAC over L2 links and on IP flow over L3 links.The support of multi-chassis LAG, meaning that a single device can connect using LACP, in a dual-home active/standby scheme to two ports in two different DARs, will be valued.

5.4 BFD

The proposed platform must support Bi-directional Forwarding Detection (BFD) for failure detection as described in RFC5880. In particular, it is required to support BFD in Ethernet interfaces, and in LACP interfaces as described in RFC7130. It is also required to support BFD in the signaling protocols listed in section 6.1, and in MPLS LSPs.

5.5 L2VPN

In the provision of L2 services for enterprise customers, the DAR shall be MEF 2.0/3.0 compliant, at least with regards to E-LAN and E-LINE services. E-TREE services will be valued. It will be very beneficial that the proposed platform is certified accordingly.
In particular, the platform must support the implementation of E-LINE services based on VLL technology, and E-LAN services based on VPLS. Support of H-VPLS topologies will also be valued.

Requirements below have already been tackled in previous sections, but they are included here as well for specific clarification of their usage in a L2VPN environment.

- The proposed platform must be able to forward traffic into the L2VPN at least based on the following encapsulations: no tagging, VLAN tagging (IEEE 802.1q), VLAN double tagging (IEEE 802.1ad) and MACinMAC (IEEE 802.1ah). For double tagging, the TPID value must also admit 0x8100 in the outer VLAN, together with the standard 0x88A8.
- The proposed platform will permit the establishment of CIR, EIR, CBS and EBS parameters over the L2VPN service and/or each specific L2VPN client connection.
- The proposed platform will permit configuring specific Layer-2 based filters per each L2VPN client connection. Other standard security measures shall be supported as well.
- For E-LAN (and E-TREE, in case they are supported) services, it must be possible to configure in the proposed platform a different rate limit per each type of BUM traffic (one for broadcast, one for unknown unicast and one for multicast). It must also be possible to configure limits in the number of learnt MAC addresses per service and client interface, and the lifetime for these learnt MAC addresses.

Finally, redundancy mechanisms, loop detection mechanisms and OAM functionalities shall be included in the proposed platform. IGMP snooping must also be supported.

5.6 L3VPN

Full-mesh and Hub & Spoke models of BGP/MPLS L3VPN must be supported for IPv4 and IPv6. In particular, the proposed platform must support the establishment of a L3VPN over multi-AS backbones.

The non-exclusive set of protocols outlined in Section 6.1 (RIP, IS-IS, OSPF and BGP-4) including static routing, must be supported in the CE/PE link. BFD associated with these protocols and DHCP Relay must also be included.

With regards to the rest of the features, like QoS, multicast, filtering, redundancy, etc., the same requirements as listed in sections 5 and 6 for IP environments must be supported, both in the client and network interfaces. Additionally, control plane control mechanisms (e.g. BGP dampening) shall be available.

5.7 E-VPN

The proposed platform must support E-VPN services. All the requirements specified above for MPLS based VPN services shall apply as well to corresponding E-VPN variants. E-VPN must be supported with single and multi-homing (with single-active and all-active modes supported for the latter case).

5.8 Quality of Service

A high-level list of QoS functionality include:

- QoS packet marking and rewriting
- Traffic Policing and Shaping
- Multi-level queuing based on priority
- PQ, WRR, WRED
- Bandwidth/Congestion Control
- Packet buffers

06 DAR Management, Programmability and Security

06 DAR Management, Programmability and Security

The following section describes requirements for the management and monitoring of DAR devices, as well as security considerations affecting the management, control and forwarding plane.

6.1 Management

As stated in section 4.4, both out-of-band and in-band management must be supported. For this, mechanisms based on SSH, CLI and Netconf will be preferred to those based on web browser.

The CLI will permit the configuration of different access profiles, at different levels and with different permissions (e.g. operator, monitoring…). It will also permit the login of multiple concurrent users.

The following protocols shall be supported:

- SNMPv2c/v3, for management and configuration purposes
- Netconf for management and configuration purposes
- NTP, for time synchronization with the rest of the network
- SCP or SFTP, for file transfer (e.g. configuration files or software upgrade versions)

6.2 Monitoring

SNMP may also be used for monitoring purposes. In that sense, the platform shall support the required MIBs for monitoring of the main activated functionalities, and it must be possible to send monitoring traps to multiple destinations.

The support of SYSLOG is also a requirement for logging redirection: all relevant system events must generate SYSLOG messages, which may be sent to multiple SYSLOG servers. Finally, for monitoring of quality of service, support of Y.1731, CFM 802.1ag, Y.1564, PNPM, IPSLA or RPM, OWAMP and TWAMP mechanisms is requested.

6.3 SDN and Programmability

As Network operators are moving away from the CLI and towards network programmability, DDBR device being a key part of the Transport network shall conform with the Open Transport Architecture defined by the MUST subgroup[4]. That network programmability can be achieved by employing a hybrid SDN hierarchical architecture, in which the management and control functionalities are split between the devices and the controller.

The main goals of such SDN solutions are:

- Agile Network Programmability, enabling full network automation and reduced time-to-market service creation.
- Network Abstraction, simplifying Operation Support Systems (OSS) and orchestrators, and their interactions, by performing the adequate level of abstraction at each layer.
- Network Intelligence, enabling Traffic Engineering (TE) and automated service provisioning mechanisms between different layers and different vendor technologies.
- Compliancy with the data models and protocols defined under the MUST specifications documents.

[4]MUST Open Transport SDN Architecture, https://cdn.mediavalet.com/usva/telecominfraproject/03V-53HVHE2_sr3_nk47_Q/WPd6tLiuS0CDkcG5S6Etug/Original/OpenTransportArchitecture-Whitepaper_TIP_Final.pdf

6.4 Network Telemetry

DAR shall support advanced monitoring and telemetry features, in particular:

- gRPC Network Management Interface (gNMI), gPB (Google Protocol Buffers) proto3 for encoding, and data exported (modelling) based on YANG models
- YANG push (RFC 8639, RFC 8640, RFC 8641, RFC 8650) with support for YANG QoS models

Those will be used by the SDN controller in order to monitor the status of the platform and the different services instantiated in the node.

6.5 Security

The following is a non-strict list of requirements having to do with security of the platform. Participating operators strongly encourage vendors to keep these capabilities up to date, incorporating new mechanisms as soon as they become available.

The proposed platform must be capable of implementing segment, packet and frame PDU filters in any physical or logical interface, and in the incoming and/or outgoing directions simultaneously with fine-grained – per subscriber - granularity. Filters will be able to take into account any combination of the following arguments: source/destination IP address, source/destination port, protocol, source/destination MAC address, VLAN, TCP flags, fragmentation flags, ICMP type and packet size. There must be counters available for each rule in the filter, increasing with each matching packet, and available for consultation via SNMP.

With regards to management operations, the platform must support the following:

- TACACS+ for authentication and authorization of the CLI features.
- Logging of all the commands executed by all the operators, for audit purposes.
- SSH sessions with 3DES encryption – and not preclude options to use more advanced encryption methods e.g. AES-128 or AES-256.
- Restriction of the management access only to a defined subset of IP addresses.

In the data plane, the platform must support Unicast Reverse Path Forwarding (URPF), as defined in RFC3704. Further, for PPP Termination and Aggregation (PTA) sessions, uRPF must be applied per subscriber. It must also be capable of limiting the maximum number of MAC addresses and limiting the BUM (broadcast, unknown unicast and multicast) traffic per physical or logical interface.

Protection mechanisms against Denial of Service (DoS) attacks is also required, at least for the following list: Tear Drop, Ping of Death, Smurf, Fraggel, UDP Flood and SYN-ACK. Mechanisms to moderate control protocol (e.g. LCP) traffic load per subscriber (requests rate, filters, etc.) shall be implemented. Finally, the proposed architecture must ensure the maximum possible isolation between the control and management planes, and the data plane.

6.5.1 Access Security and Anti-Theft

In general, the solution must support the necessary security mechanisms to authenticate and encrypt communications between the network element and its management system or controller.

The DAR must offer the possibility of enabling this traffic only after it has been authenticated by the management platform/controller.
The system should also offer the possibility to enable anti-theft mechanisms that prevent the use of the equipment in any other environment than the one it was conceived for.

07 Additional Requirements

07 Additional Requirements

To finish this specification, the following sections include additional requirements which do not fit any of the software packages defined above, but that are mandatory independently on the DAR flavor.

7.1 Configuration and Versions Management

It shall be possible to modify the node configuration by means of configuration files accessible or directly copied in the node disk (via SFTP or SCP). Processes of file transfer, loading of a target configuration, and activation of the target configuration shall be three different processes. None of this shall require a reboot of the node, nor will affect its operational status.

It will be valued that more than two (active and target) configuration versions are accessible to the operator. This will be achieved by storing more than one target configuration in the node. It shall be possible to activate any target configuration from those stored.

Prior to the activation of a target configuration, the device will check its coherence, and will warn the operator of any incompatibility.

For configuration changes based on CLI:

- No configuration change will require a reboot of the system for its activation
- It will be valued that the introduction of the command, and its activation, are two different processes (e.g. using “save configuration” or “commit” commands).

On the other hand, notwithstanding automated procedures for Zero-Touch Provisioning as described in section 10.3, it shall also be possible to transfer a software package to the node for a version update or downgrade. Again, the process of transferring the file, and activating it, must be two different processes, and none must require a reboot or affect its operational status. Indeed, it is strongly encouraged that, even for those proposals that are built using a pizza-box approach, and do not implement redundancy of the control board, ISSU is supported.

7.2 Zero-Touch Provisioning

Zero-Touch Provisioning (ZTP) is the process to deploy a Network Operating System (NOS) and a base configuration in a network element, so it can enter in production environment without any human configuration. ZTP process is executed for the first time when the device is turned on in the network.

Periodically, system vendors release new versions of their NOS; a very similar process can be done for upgrade scenarios. Nowadays, the NOS upgrade process is a vendor-dependent process, which differs between each vendor solution. The ZTP and upgrade mechanisms should evolve to generate a common procedure for open white-box scenarios.

This procedure will influence the requirements with regards to configuration management. The ZTP platform or the SDN controller will have to maintain the NOS images and a pointer for the current configuration files that the network element requires.

7.3 Licensing

It is not the goal to define a strict licensing agreement at this stage, but some important ideas, in line with the disaggregation philosophy, would include:

- With regards to SW packages defined in sections 5 to 8, it should be possible to differentiate between licenses associated with each of these packages. The figure of an “all included” license should also exist.
- Upgrades to features existing in one of the packages should not affect the cost of other SW packages.
- The platform shall be compatible with a pay-as-you-grow model not only from the hardware point of view, but also as defined by software licenses.

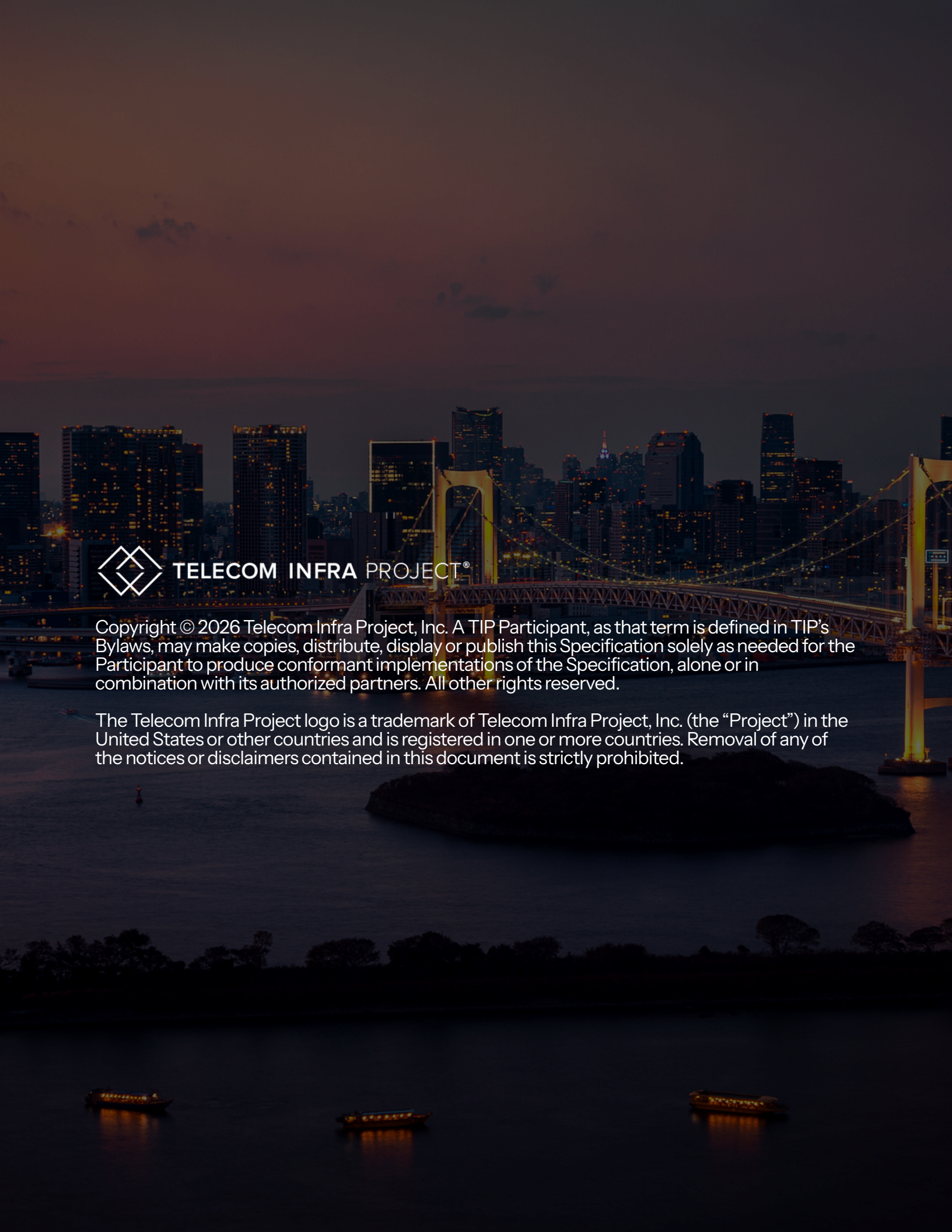
08 Glossary

08 Glossary

AC	Alternating Current
API	Application Programming Interface
AS	Application Programming Interface Autonomous System
ASIC	Application-Specific Integrated Circuit
BFD	Bidirectional Forwarding Detection
BGP-4	Border Gateway Protocol 4
BGP-LS	BGP Link State
BGP-LU	BGP Labelled Unicast
BMCA	Best Master Clock Algorithm
CBS	Committed Burst Size
CE	Customer Edge
CIR	Committed Information Rate
CLI	Command Line Interface
CoS	Class of Service
CPU	Central Processing Unit
DC	Direct Current
DSCP	DiffServ Code Point
E-VPN	Ethernet VPN
EBS	Excess Burst Size
ECMP	Equal Cost Multi Path
EIR	Excess Information Rate
ER	Extended Reach (optics)
EXP	MPLS EXPerimental bits
FIB	Forwarding Information Base
FPGA	Field-Programmable Gate Array
FRR	Fast Reroute
gNMI	gRPC Network Management Interface
GPS	Global Positioning System
GRE	Generic Routing Encapsulation
gRPC	gRPC Remote Procedure Call
H-QoS	Hierarchical Quality of Service
ICMP	Internet Control Message Protocol
IP	Internet Protocol
IS-IS	Intermediate System to Intermediate System
L2TP	Layer 2 Tunneling Protocol
L2VPN	Layer 2 Virtual Private Network
L3VPN	Layer 3 Virtual Private Network
LACP	Link Aggregation Control Protocol
LAG	Link Aggregation Group
LAN	Local Area Network
LDP	Label Distribution Protocol
LEA	Law Enforcement Agency
LED	Light Emitting Diode
LFA	Loop Free Alternate
MAC	Media Access Control
MBH	Mobile Backhaul

MPLS	Multiprotocol Label Switching
MPLS-TP	MPLS Transport Profile
NBI	North-Bound Interface
NIC	Network Interface Card
NNI	Network-Network Interface
NOS	Network Operating System
NTP	Network Time Protocol
OAM	Operations, Administration and Management
OCP	Open Compute Project
OLT	Optical Line Termination
ONIE	Open Network Install Environment
OSPF	Open Shortest Path First
OWAMP	One-Way Active Measurement Protocol
PBB	Provider Backbone Bridge
PCEP	Path Computation Element Protocol
PE	Provider Edge
PIC	(BGP) Prefix Independent Convergence
PIR	Peak Information Rate
PNPM	Passive Network Performance Monitoring
PTP	Precision Time Protocol
PW	Pseudowire
PWE3	PW Emulation Edge to Edge
QinQ	802.1Q Tunnelling
QoS	Quality of Service
RED	Random Early Detection
RFC	Request For Comments
RIP	Routing Information Protocol
RSVP-TE	Reservation Protocol – Traffic Engineering
SC	Standard Configuration
SCP	Secure Copy Protocol
SDN	Software Defined Networks
SFP	Small Form-factor Pluggable
SR	Short Reach (optics)
SKU	Stock Keeping Unit
SLA	Service Level Agreement
SME	Small or Medium Enterprise
SNMP	Simple Network Management Protocol
SSH	Secure Shell
STB	Set-Top Box
T-BC	Telecom Boundary Clock
T-LDP	Targeted LDP
TACACS	Terminal Access Controller Access Control System
TCP	Transport Control Protocol
TI-LFA	Topology Independent LFA
TLS	Transport Layer Security
TWAMP	Two-Way Active Measurement Protocol
UDP	User Datagram Protocol
UNI	User-Network Interface
UPF	User Plane Function
URPF	Unicast Reverse Path Forwarding
USB	Universal Serial Bus
VLAN	Virtual Local Area Network
VLL	Virtual Leased Line
VM	Virtual Machine

VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VRF	VPN Routing and Forwarding
VRRP	Virtual Router Redundancy Protocol
VxLAN	Virtual Extensible Local Area Network
WAN	Wide Area Network
WFQ	Weighted Fair Queueing
WRR	Weighted Round Robin
ZR	Extended Reach (optics)
ZTP	Zero-Touch Provisioning



TELECOM INFRA PROJECT®