Telco Edge Cloud: Edge Service Description and Commercial Principles

Whitepaper

October 2020

About the GSMA

The GSMA represents the interests of mobile operators worldwide, uniting more than 750 operators with over 350 companies in the broader mobile ecosystem, including handset and device makers, software companies, equipment providers and internet companies, as well as organisations in adjacent industry sectors. The GSMA also produces the industry-leading MWC events held annually in Barcelona, Los Angeles and Shanghai, as well as the Mobile 360 Series of regional conferences.

For more information, please visit the GSMA corporate website at www.gsma.com. Follow the GSMA on Twitter: @GSMA.

About the Telco Edge Cloud Group

The Telco Edge Cloud (TEC) group brings together over 20 operators, covering all regions, who are working to promote a collaborative deployment of cloud capabilities at the edge of their networks. TEC is aiming to align Multi-Access Edge Computing (MEC) business models, charging principles and commercial deployment considerations.

TEC is working in partnership with the GSMA Operator Platform Group, which aims to create the architecture and technical requirements to guide other Standard Developing Organisations (SDOs) in the development of specifications.

Executive Summary

The Telco Edge Cloud taskforce was launched in the GSMA in March 2020 by 19 operators with the intention to design and develop a global edge computing service based on the federation of the edge infrastructures and platforms of a set of operators and edge service providers. Following its principle of being open and inclusive, the taskforce has grown to 25 members at the time of edition of this whitepaper and has set the basis for the global service launch. This document includes a first description of the basic services the Telco Edge Cloud (TEC) will provide along with a set of commercial principles around stakeholder relationships and charging schemes.

This whitepaper is complemented by other documents produced by the GSMA, like the Operator Platform Telco Edge Proposal, which describes the technical framework and requirements of the platform that will support the Telco Edge Cloud service.

The document will help:

- Operators and other service providers:
    - assess the revenue and business potential in edge computing;
    - analyse technological alternatives;
    - consider the impact of TEC on their portfolios; and
    - Ultimately select a role to play in this new ecosystem
- Operator Platform group (OPG) complete the specification of requirements, incorporating those needed to accommodate the TEC commercial principles;
- Technology providers to identify potential ways to contribute to TEC (along with the analysis of OPG technical details);
- Future TEC customers understand the kind of services TEC will deliver and the value they may bring to them and their end users.

1. Introduction

1.1 Overview

The Edge Cloud inherits most principles, mechanisms, and tools from the traditional cloud, but provides Customers with additional benefits as the edge brings proximity to the end user devices, enabling characteristics like lower latency, relevant for time-sensitive applications, and transport efficiency, for data intensive ones. The Operator Platform Group (OPG) has defined the Edge Cloud functionality in the OPG.01 Operator Platform Telco Edge Proposal and the Telco Edge Cloud (TEC) taskforce works on launching a global telco edge computing service that implements it.

The Telco Edge Cloud provides other advantages like trust, security, mobility support, privacy or managed connectivity based on its optimal integration with the network systems, that allows to combine computing and connectivity performance with a secure network environment.

Operators that participate in TEC decided to perform Edge Cloud trials over live networks with Customers and end users, and subsequently update the supporting standards based on their experience. These trials will take place in the course of 2020/21 in an organised way and with the appropriate agreements in place. A basic requirement is to have a service description and commercial requirements described in a GSMA official document which can be duly updated over time without the need to touch any signed agreements.

1.2 Scope

This document contains a service description of the expected functionalities of the Telco Edge Cloud service. It further includes the commercial principles for the relationship between the operators that participate, as well as general guidelines about what needs to be agreed upon with Customers and technical providers.

2. Description of Telco Edge Cloud

The TEC initiative was started by operators following the Operator Platform concept developed within the GSMA, with the goal to monetise their network capabilities, including their edge computing resources, in a standardised and consistent way and as such serve the global demand for innovative, distributed and low latency services. TEC allows application developers, service providers and enterprise customers to develop and deploy their solutions over a global telco footprint from a single and simple interface, enjoying all the benefits from a secure and reliable distributed telco infrastructure. Furthermore, by providing a common interface and engagement model across disparate Operator networks, the TEC facilitates the development of mutually beneficial, synergistic, and sustainable partnership opportunities between cloud service providers and the Operator community.

The Telco Edge Cloud is created by the federation of TEC Service Providers that interconnect their TEC platforms. A TEC platform provides the TEC Services described in this document and follows the Operator Platform's technical reference [OPG.01].

The TEC network is the result of this interconnection of TEC platforms and enables the delivery of global Edge Computing services across the federation footprint. The interconnection is performed

using standard federation interfaces, directly between platforms or through a TEC Federation Hub, offering any of the federated TEC platforms access to the entire TEC network.

The TEC is a highly distributed cloud computing system integrated with the telecom network that offers traditional cloud computing services (IaaS, NaaS, PaaS, SaaS) with higher proximity to End Users. It further enhances traditional cloud computing services with features that improve and guarantee a certain application performance at the edge.

The TEC follows the principles and recommendations described in the GSMA [OPG.01] for federated edge computing platforms.

3. Telco Edge Cloud services description

This section describes the services the Telco Edge Cloud will provide.

The TEC follows the Cloud Computing paradigm: enable network access to a scalable and elastic pool of shareable physical and virtual resources with self-service provisioning and administration on-demand.

The TEC services will range from basic infrastructure as a service (IaaS) to more advanced ones (NaaS, PaaS, SaaS). As in the current public cloud, “as a Service” means delivered “over the internet” through a self-service portal or programmable APIs and, thus, will require the edge cloud to be embedded in the network but connected to the public internet as well.

The TEC services may provide the same service categories as traditional cloud:

Infrastructure as a Service (IaaS): Cloud computing service in which the capabilities that are provided to the cloud service Customer are processing, storage, and networking resources:

- The IaaS Service Provider (SP) delivers virtualised infrastructure capabilities over the internet to the Customer. This includes both physical data centre resources such as servers, storage and networking hardware, and the virtualisation layer, consisting of hypervisors (VMs for computing, Software Defined Storage for storage and Software Defined Networks (DC-SDN) for networking) or Container Infrastructure Service Managers (CISMs).

- The IaaS SP deploys and maintains the physical Infrastructure and the virtualisation layer, and provides additional services like billing, monitoring, identity and access management (IAM), security features, load balancing and clustering, HA-related backup, replication and recovery

- The IaaS Customer is responsible for installing, maintaining, and managing the software layers (OS, middleware, data, and applications). Typical IaaS operations include:
  - create Virtual Machines (VM) or Container Clusters.
  - install Operating Systems (OS)
  - deploy middleware (e.g. a database (DB))

- create or trigger backup processes
- install an application into the VM or container
- manage the system: track usage and performance, troubleshoot app issues, recover from disasters, balance network traffic, etc.
- define app policies

Platform as a Service (PaaS): Cloud computing service in which the capability that is provided to the cloud service Customer is the functionality to deploy, manage, and run Customer-created or Customer-acquired applications using one or more programming languages and one or more execution environments supported by the cloud service provider.

- Builds on top of the IaaS service. In addition to the underlying infrastructure components, providers offer, host, and manage operating systems, middleware, and other runtimes for Customers.

PaaS simplifies workload deployment, but also somewhat limits Customer's flexibility to create the environments they want.

Software as a Service (SaaS): As per NIST definition [NIST 800-145]: SaaS applications in the cloud are made accessible via a network to the SaaS consumers. The Customers of SaaS can be organisations that provide their members with access to software applications, end users who directly use software applications, or software application administrators who configure applications for end users. The Customer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. All this complexity is managed by the SaaS provider.

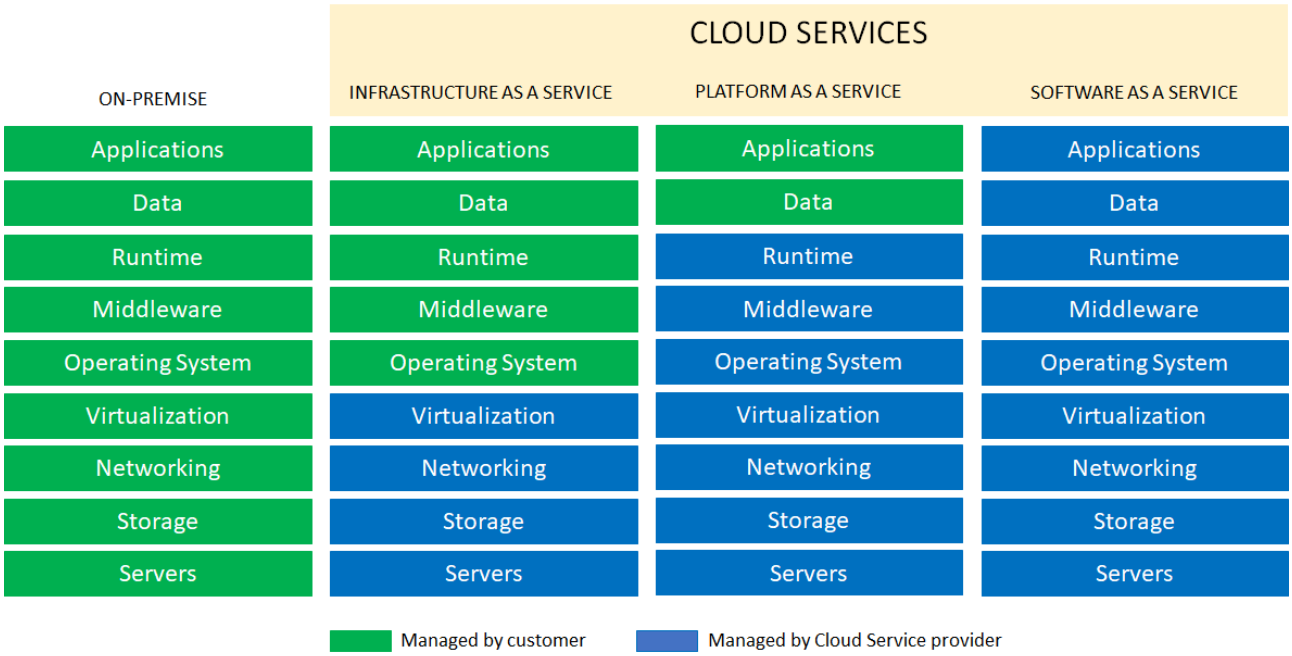

Figure 1 – Cloud Services

Network as a Service (NaaS): As defined in [ITU-T Y.3515] Cloud computing service in which the capability provided to the cloud service Customer is transport connectivity and related network capabilities. The NaaS functional architecture aims for a scalable and programmable on-demand

network allowing the NaaS Customer to provision network services and resources, as needed, automatically or with minimal interaction with the NaaS provider. In addition to the cloud computing characteristics described in the paragraph below, NaaS usually includes programmability, the ability to access services features through application programming interfaces (APIs). Examples include programmable quality of service (QoS) and network policy rules.

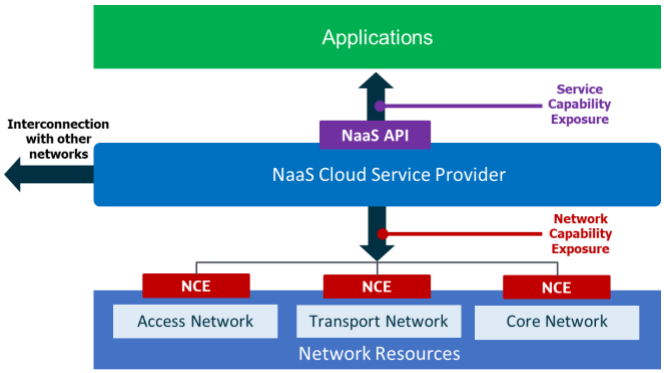

Figure 2 – NaaS Services

The Telco Edge Cloud will mainly focus on IaaS and PaaS, offering NaaS as it becomes developed and standardised. SaaS is not in the scope of TEC but, in case it becomes in the future part of its portfolio, it is likely to happen using TEC partners.

3.1 Telco Edge Cloud Service Characteristics

The same characteristics defined in [ITU-T Y.3500] for cloud computing services apply to Telco Edge Cloud as well:

- Broad network access: the edge computing infrastructure can be connected to fixed or mobile networks: 5G/LTE, FTTH, LAN, WiFi, etc.
- Measured usage: usage can be monitored, controlled, reported, and billed in order to optimise and validate delivered services, and bill Customers according to what they ordered or used.
- Multi-tenancy: physical and virtual resources may be allocated in a way that workloads or resources of multiple tenants are isolated from one another.
- Self-service: ability to provision cloud capabilities, as needed, automatically or with minimal interaction with the cloud service provider. Self-services such as self-provisioning, self-care and self-design reduce the costs, time and effort needed to take an action.
- On-demand: ability to deploy, migrate, adjust, and maintain services rapidly and as needed that reduce time-to-market and accelerate innovation.
- Rapid elasticity and scalability: physical and virtual resources can be rapidly and elastically adjusted, in some cases automatically, to quickly increase or decrease capacity (scale out, in, up or down) in response to a change in usage demand. For Customers, resources available for usage will appear to be unlimited and they do not need to worry about resource limitations and capacity planning. In the case of edge cloud, this characteristic can be more difficult to be achieved economically due to the smaller resource pools and lower

multitenancy ratios. In general, a trade-off will have to be made at the edge between elasticity/scalability and cost-efficiency.

- Resource pooling: cloud physical and virtual resources may be aggregated in order to serve one or more Customers while abstracting the complexity of the process from the Customers. Customers will be able to specify location at a higher level of resource pool (e.g. country, state, or a data centre).

In addition to these, TEC offers additional characteristics as described below.

3.1.1 TEC Infrastructure as a Service (IaaS)

Telco Edge Cloud adds several key features to the traditional IaaS service described above:

- Intent-based multiple virtual infrastructure creation: ability to create VMs or containers in multiple locations at different operator domains based on certain requirements (geographical area to cover, latency, residency, bandwidth etc.) with a single order.

- Notification of SLA evolution and breach, so that the Customer can take actions (scale up/down, change configuration, etc) if any of his requirements cannot be met.

3.1.2 TEC Edge Connectivity as a Service (NaaS)

TEC will provide NaaS capabilities to manage the connectivity between users and the edge computing platform, between components of the platform and among different platforms, for:

- **Dynamic adaptation of network performance or quality to a Customer/application demand**; As an example, a real-time game session, a remote video production event, or a remote surgery session may require specific network parameters (upstream/downstream bandwidth, latency, jitter, etc.) for certain Customers at a certain point in time. NaaS will allocate the required network resources to provide certain performance or quality parameters to the Customers.

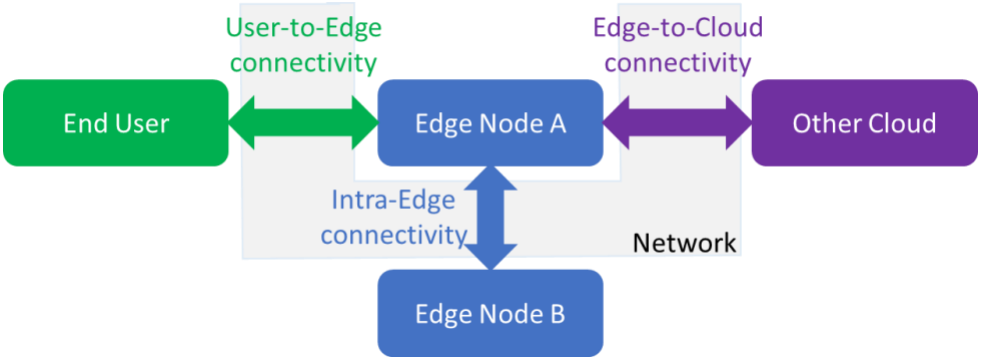

Figure 3 – TEC Edge Connectivity as a Service

Overall, the network connectivity at the edge needs to be adapted to certain fluctuations:

- The **scalability of virtual compute and storage resources** is coordinated with the corresponding **scalability of the networking resources connecting them**. As VMs and containers scale up and down, the corresponding networking environment will adapt dynamically and automatically to those changes so that overall performance, reliability, and quality of the Operator Platform deployed applications are assured.

- **On-demand (re)configuration**; as computing and storage resources are reconfigured or migrated, or users move from one edge zone to another, on-demand dynamic reconfiguration of the corresponding networking resources will allow to satisfy the requirements of the edge computing apps wherever they are moved.

- **Adaptation to heterogeneous network technologies**; the distributed edge computing systems may be connected using different network technologies.

Some of the features that TEC should provide for this NaaS service are:

- Administration:

    - Quality of Service profile selection: TEC Service Providers offer Customers a list of predefined QoS profiles to choose from to meet the enhanced connectivity requirements of a certain edge application Customers want to deliver. The catalogue is dynamic, and the TEC Service Provider may enhance it progressively as new categories of edge applications are identified that pose specific connectivity requirements. Customers can monitor the received QoS by accessing a set of predefined KPIs.

    - Policy setting: Customers may set policies for the use of the enhanced connectivity associated to certain applications. Such policies may define conditions for provisioning and use of that enhanced connectivity. The TEC Service Provider delivers the logic for controlling and applying the policies when End Users use the applications.

- Provisioning and use:

    - The TEC Service Provider dynamically configures connectivity as End Users connect following the QoS profiles and policies that the Customer has set, whenever this is feasible taking into consideration possible conflicts with other services. The TEC Customer is notified about the result of the connectivity provisioning that may include the recommendation of a feasible alternative QoS profile or a rejection of the connectivity request.

    - If the enhanced connectivity should be delivered on the End User demand, it will only be established when the set policy applies, for example an explicit End User request by pressing an “activation” button. Once the enhanced connectivity is set up, the End User is informed, and the usage accounting information is generated for the Customer.

3.1.3 TEC Platform as a Service (PaaS)

As any other PaaS, TEC PaaS includes application enablement services, and allows Customers to focus on the data and application logic and abstract from the physical infrastructure, virtualisation layer, operating system, middleware, and runtime environment.

The TEC PaaS is based on Containerised infrastructure, as for example Kubernetes. Containers are becoming the most common virtualisation mechanism and a must for cloud-native applications.

TEC PaaS also includes some extended functionalities, like the identification of the optimal edge node to deliver an application to a certain end user, 3GPP system integration or interconnection of different edge platforms, that:

- allow operators to deliver a differentiated service,
- facilitate resource federation among global TEC participants, and
- enables the provision of edge services when End Users move from one edge zone to another, including roaming beyond their home network.

The TEC platform provides the TEC Customer the capability to:

- On-board/register an application (in an application repository).
- Instantiate an application in edge nodes anywhere in the TEC across multiple operators and edge service providers in different modalities:

    1. Static application instantiation: The Customer must specify:
        i. Footprint/coverage area selection.
        ii. Customer reach/MNO selection.
        iii. Infrastructure resources required by the application (per deployed instance/ cluster): CPU, memory, storage, networking, specialised hardware (GPU, NPU, FPGAs, other accelerators…). For the sake of manageability, the Customer will have to choose from a set of predefined resource profiles.
        iv. Specific functions consumed by the application (e.g. databases).
    2. Dynamic application instantiation: in addition to the parameters above, the Customer can set rules for instantiating an application in a cloudlet (edge node) based on:
        i. User demand: like the amount of requested edge compute resources or the number of users.
        ii. Application performance requirements: latency, jitter, bandwidth, etc.
        iii. Privacy requirements: specific geographical area.
        iv. The dynamic application instantiation is applied also to serve users roaming in a visited network.
        v. The number of application instances and edge computing resources assigned to a TEC Customer scales up and down automatically as application usage increases or decreases, respectively.

- Use the right edge node to serve an edge application to a certain end-user/device at a certain location:

    3. the selection of the serving edge node is done by the TEC platform and can be based on Customer or network policies, like optimal use of network or computing resources or meeting specific performance requirements (latency, bandwidth, proximity, etc).
    4. The TEC platform may:

        i. change the application instance that serves a certain end user to another one in a different edge node as the user moves or the original serving edge node performance degrades, so as to keep the required application performance parameters,

        ii. move the application instance serving a certain end user to a new edge node following a change of user location (mobility) or a performance degradation (this applies for application instances serving a single user)

- Upscale/downscale automatically application instances as the number of users attached to them grows or decreases.

3.2 Uniform Service Experience

For TEC to appear as a consistent service to TEC Service Customers and allow extending the reach of the service offering beyond their own footprint, the TEC Service Providers should offer a homogeneous service. That means that they should align to ensure a consistent approach in their definitions of:

- The Customer interface (NorthBound Interface as in OPG.01). The TEC Customer has access to TEC services and functionality across the whole TEC footprint by using a single interface, the one of any of the federated TEC platforms.

    - The Customer should be able to change the TEC Service Provider and move to another platform in the TEC network without requiring any change in the application code or in the management and administration routines.

- The granularity of the regions in which the service is offered (e.g. city level, province level, county, state, etc.)

    - Potentially multiple levels of granularity could be offered to address different use cases requiring edge compute

- The set of IaaS and PaaS resource types offered

    - Alignment should be on a minimum set of resource profiles or flavours (e.g. well-defined resource types for storage intensive activities, another types to support compute intensive workloads, specific ones for network intensive functions, etc).

    - If offering multiple levels of granularity for the regions, alignment could be on a common resource set for each level of granularity.

3.3 Management and Administration Functionality

The TEC Service Provider is able to perform the following tasks following uniform TEC mechanisms (those prescribed by the Operator Platform Group):

- Administer resources:
  - On-board/register infrastructure resources (computing, storage, networking).
  - Define, instantiate, and activate regions/areas.
  - Assign resources to those regions/areas.
  - Expose those region/areas and resources in a catalogue, so that they are made visible and accessible by TEC Customers.
  - Consult regions/areas in other federated TEC platforms,
- Monitor usage, faults, alarms, and performance of the edge resources.
- Issue and manage trouble tickets in both directions: platform-operator, operator-platform.
- Set quotas, lock resources
- Track service logs
- Account and charge for the services, based on transactions and/or resource usage, with pricing that may vary based on region, time, and demand
- Access and use a portal for management and administration routines, like consulting information such as the end user subscription or device location.
- Transfer of a customer's service request to a federated TEC Service Provider when associated to a region/area out of his footprint, or when the request explicitly states it must be addressed to that Service Provider.

The Customer is able to perform the following tasks using uniform TEC interfaces (those prescribed by the Operator Platform Group)::

- Access available regions (areas covered by edge nodes) and resources across the global Telco Edge Cloud
- Monitor their resources across the entire Edge Cloud, including application instances, events and alarms, and performance. The TEC Service Provider can deliver this monitoring via dashboards or via APIs that can feed the customer's own dashboards.
- Set rules and policies to instantiate apps, allocate resources and receive event notifications
- Benefit from a clear pricing scheme for each region/operator/instance/service
- Access and use APIs and portals for management and administration routines

The TEC will provide generic management and administration features:

- IAM, providing the credentials for the TEC Customer to access the management and administration functionality.
- End User access control: to define the list of end users authorised to use a certain service or application.
- Order management
- Reporting and Monitoring
- Event, fault and alarm management

- **Accounting** and **charging** mechanisms, based on transactions (e.g. app onboarding/ instantiation requests), resource allocation (e.g. bandwidth, latency) and/or resource usage (data volume used).
    - Pricing may be modulated based on region/area, demand, time/day, operator…
- **Security** framework (authorisation/authentication/encryption/firewall/attack prevention), including verification of security for onboarded applications.
- **Catalogue** management including regions, services, resources, and applications.
- **Marketplace** management including onboarding, integration with ordering, billing, and payments.

3.4 Other TEC services

3.4.1 DevOps support

The TEC Service Provider could optionally provide an environment for continuous integration, continuous deployment, continuous testing, continuous verification (CI/CD/CT/CV), but it will allow the Customer to use its own environment.

3.4.2 Life Cycle Management tools for Infrastructure and Applications

, and widely used LCM cloud tools like Terraform, Ansible or Helm, with the intention to offer customers a familiar development and deployment environment

3.4.3 Functions as a Service (SaaS)

The TEC Service Provider could optionally provide SaaS services, like libraries of basic functionality like databases (DBaaS), video processing (VPaaS), AI/ML processing (AIaaS, MLaaS), etc. The set of libraries and functionalities will have to be agreed by the TEC community to be considered a TEC Service. The TEC community will have a procedure in place to take such decisions.

4. Stakeholders – Roles in the value chain

4.1 Stakeholders

Key TEC stakeholders are outlined below.

- Enterprises – customers of TEC who may already use public cloud, multi-cloud, hybrid cloud, or on-premise solutions. Some of their workloads and micro services may perform better or need the edge for technical or economic reasons. Enterprises may have complex requirements for fixed and mobile communications, IT, IoT, cloud and applications. These are often served by bespoke, outsourced, and managed solutions from Service Providers, and includes organisations from industrial, healthcare, media and entertainment (e.g. gaming) sectors. Enterprise applications may serve their own end users and devices, or those of other businesses or consumers.
- Service Providers – include system integrators, resellers, IT, operations technology and IoT solution providers, mobile/fixed operator enterprise business units and cloud hyperscalers. Service Providers often have industry vertical expertise, platforms, managed services, apps and enterprise customer base. TEC enables Service Providers to extend their reach to the edge cloud combined with unique networking capabilities and offer Customer vertical or consumer applications.
- Edge Service Providers - provide commercial edge cloud services (potentially IaaS, PaaS, SaaS, and a marketplace) to market aggregating edge resources from multiple network operators or other infrastructure providers and exposing to Customers via its own platform.
- Edge Interconnect Hub Providers - facilitate the interconnection among Operators (as Service Providers) and Edge Service Providers.
- Middleware providers – companies with own software platforms, engines, SDKs and developer community - TEC may enhance the performance of specific software platforms, engines and SDKs, thereby providing additional value to their developers.
- Independent Software Vendors and Developers – ISVs sell applications & SaaS to enterprises and consumers. Some will specialise in solutions that are enhanced or enabled by the Telco Edge Cloud.
- Server and Storage Vendors - view telco as an attractive segment and even more so as telcos invest in edge computing infrastructure. Blueprints may be needed to help drive scale and cost efficiencies. They also have strong enterprise sales channels which may support the TEC.
- Network operators – own mobile and fixed access and core network. TEC enables network operators to expose and manage their infrastructure and capabilities to market in a consistent way, thereby driving uptake and usage.

Note: One stakeholder may play multiple roles in the value chain.

4.2 Value Chain Description

Figure 4 illustrates a generic cloud value chain encompassing connectivity, facilities, data centre, cloud stack, app enablement, app delivery, business, and end-user.

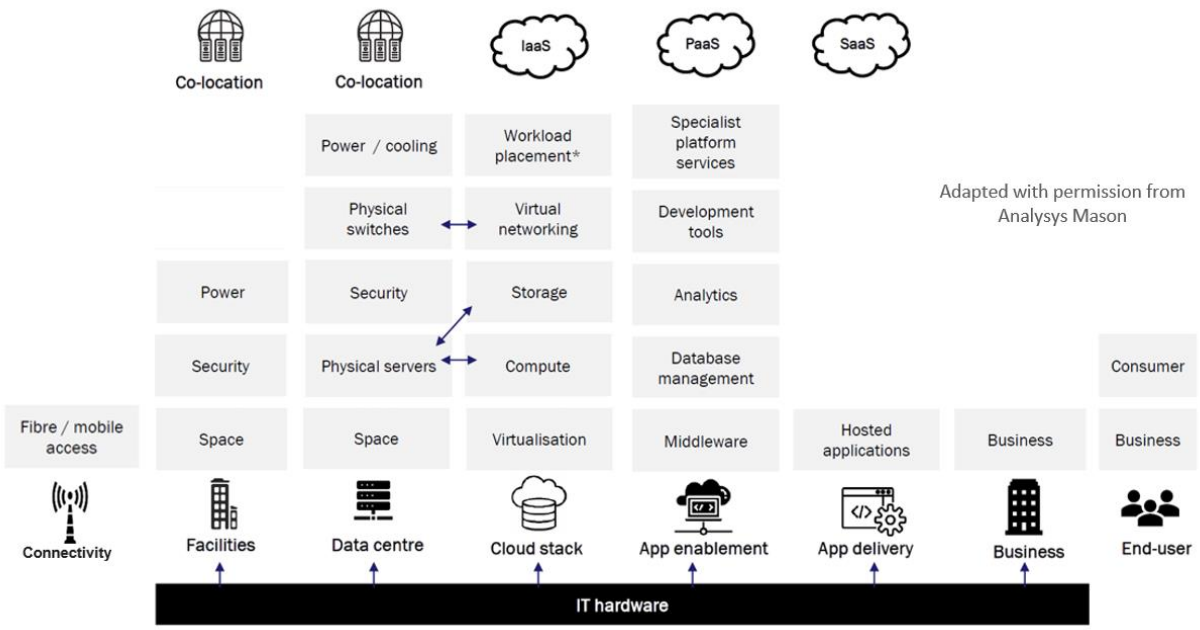

Figure 4 – Cloud Value Chain (source: Analysys Mason, Public edge cloud market sizing study)

Figure 5 illustrates two potential operator roles in the edge cloud value chain. Other roles may be possible, but this document focuses on operator as Infrastructure Provider and operator as a Service Provider.

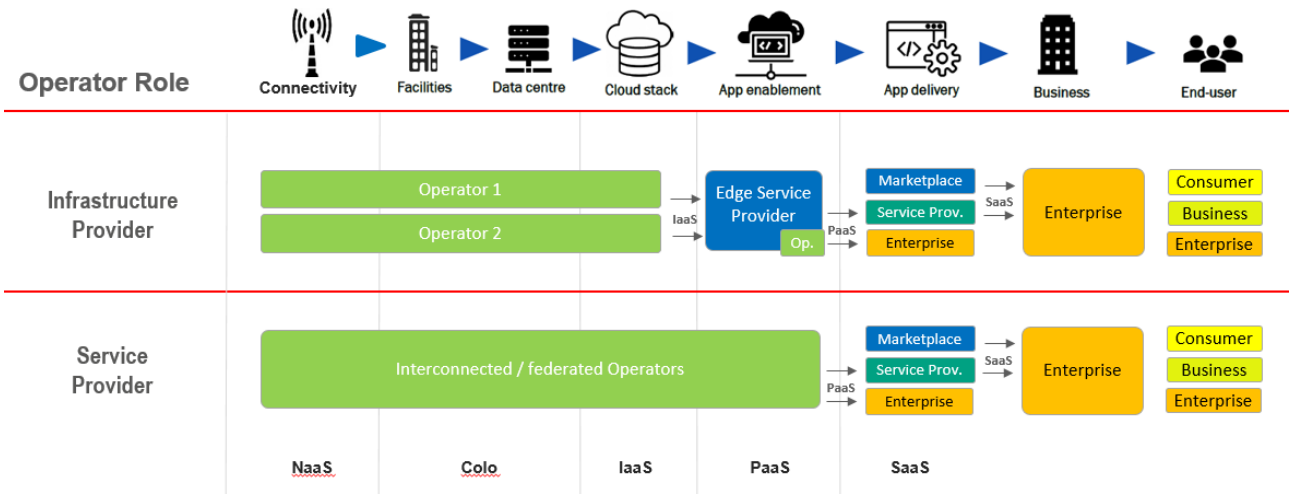

Figure 5 – Potential Operator Roles in Edge Cloud Value Chain

Operator as Infrastructure Provider – Operator provides the edge connectivity, computing, and storage infrastructure. An Edge Service Provider may use this edge infrastructure and provide aggregated edge platform services to the market and may also provide a marketplace of applications to Enterprises and Service Providers. Operators could also be a sales channel of the Edge Service Provider's services, selling solutions direct to the Enterprises.

Operator as Service Provider - Operator provides the edge connectivity, infrastructure, and edge platform services directly to the market. It may interconnect with other Operators and Edge Service Providers to provide access to edge services beyond their own infrastructure and network footprint. Operators acting together offer a federated edge cloud service direct to Enterprises and potentially to other Service Providers.

4.2.1 Relationship between Operators (as Infrastructure Providers) and Edge Service Providers

To enable an Edge Service Provider to offer edge cloud services to the market, Operators may expose their edge resources, as defined in [OPG.01], as a wholesale service to the Edge Service Provider for **aggregation**. The Edge Service Provider operates its own platform which may include software components running on Operator edge infrastructure. Two potential scenarios are illustrated in figure 6.

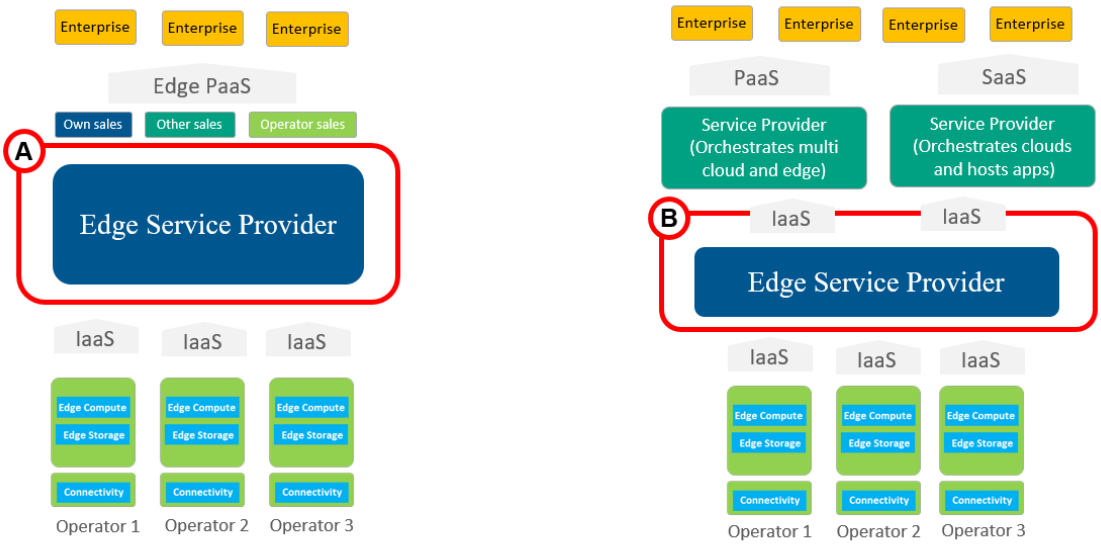

Figure 6 – Potential Scenarios for Edge Service Provider

A. Edge Service Provider offers a full commercial TEC PaaS to the market, by aggregating and exposing multiple Operator edge infrastructures via its own platform. They may also develop an ecosystem around their API/SDK and either sell service direct to Enterprises or through other channels such as Operator business units.

B. Edge Service Provider offers access to Operator edge infrastructure to enable other parties (eg Service Providers) to integrate Operator edge infrastructure into their own cloud services provided to Enterprises. The scope may include: technical resource exposure, click through agreements, service provisioning tools, usage accounting, service and fault reporting, billing, charging, etc. Their role may be a broker / facilitator or could be a wholesale service provider. If the Edge Service Provider is acting as a broker, it enables agreements between Service Providers and Operators, while providing facilitation services to both parties.

These arrangements allow enterprises to deploy their workloads on Operator infrastructure, leveraging the commercial framework and various agreements in place.

4.2.2 Relationship between Operators (as Service Providers) and Edge Service Provider

Operators which deploy their own edge platform and provide services direct to Enterprise Customers, may **interconnect** (technically via Operator Platform EWBI and potentially via hubs) to other Operators with similar capabilities and with Edge Service Providers, to extend the reach they are able to offer their customers. Together these interconnections enable the federated Telco Edge Cloud and grant an Enterprise access to edge resources across the whole federation footprint from any one of the Operator Platform providers or Edge Service Providers.

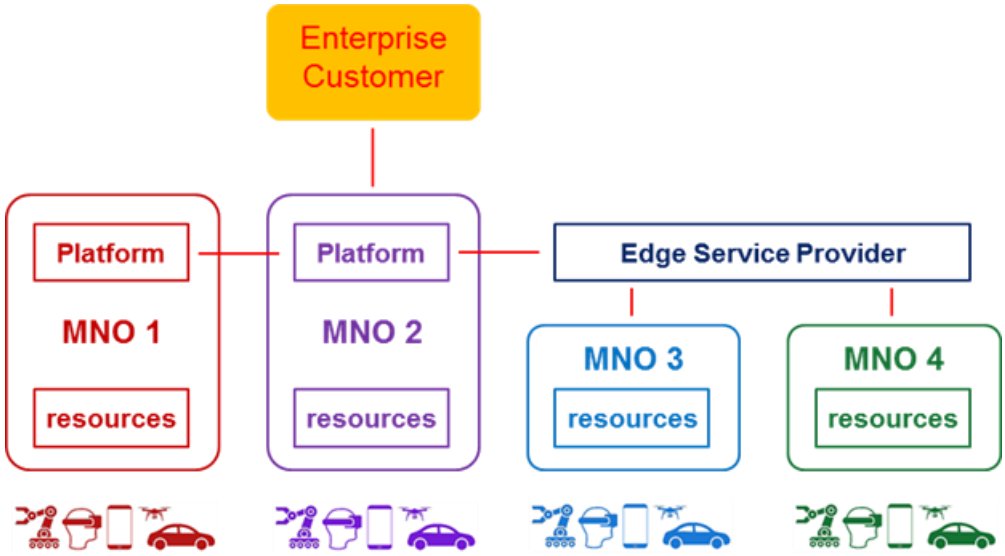

Figure 7 – Aggregation and interconnection options for federation of operator's edge resources

To facilitate the interconnection among Operators (as Service Providers) and Edge Service Providers and the creation and evolution of the global Telco Edge Cloud, one of the stakeholders may become an Edge Interconnection Hub Provider and implement a Federation Broker function (as defined in OPG.01) in order to enable one to many interconnections.

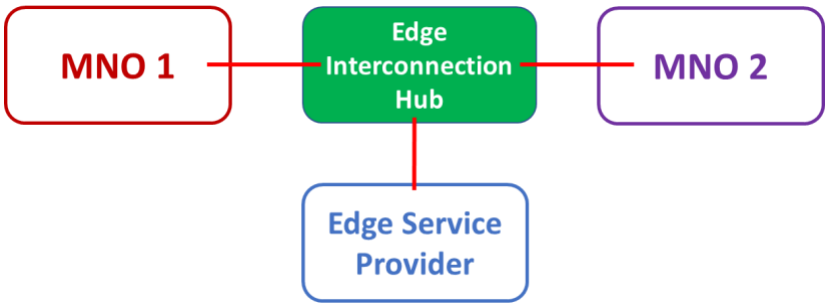

Figure 8 –Interconnection using an Edge Interconnection Hub

5. Charging Principles

With TEC, Operators provide additional value to Customers, such as reduced latency, improved speed, security, and privacy in the market which is currently partially served by the Hyperscalers.

TEC will offer familiar, widely accepted charging model options, based on Customers' needs. The charging models described below are applicable for the transactions between Service Providers and Operators (in case the SP is not an Operator), as well as between Operators. It is highly recommended to use simple models (e.g. subscription or tiered models) to ensure a smooth and quick go-to-market for the TEC offer and facilitate inter-operator charging.

5.1 Pay per Use model

Pay for resources when they are used. Resources are allocated and released dynamically following the Customer request, being served on demand. Tariff scheme depends on the type of resources that is requested. Examples include:

- Computation: based on computation capacity (number of VMs, containers or other virtual resource) and usage time. Unit price will depend on the type of server used. Several server types, with different hardware configurations (e.g. CPUs, specific processors GPU, NPU, etc), will be available for the Customer to choose with different tariffs. The choice usually will depend on the kind of load the Customer needs to execute.
- Storage: based on volume of data stored (GB) and usage time. Storage is dynamically scaled up or down as demanded by the Customer. Tariffs differ depending on the storage type requested (e.g. SSD/HDD).
- Connectivity: Amount of data transferred (e.g. per GB) upstream (to the edge) or downstream (from the edge). Tariff may vary depending on the connectivity requirements: speed (guaranteed bandwidth), latency or direction (upstream vs downstream).
- Edge Connectivity as a Service: based on connection type requested and usage time. Connection types may include, among others: device to edge, edge to edge (between cloudlets) and edge to cloud connections.

There might be charges per API request (transaction) or per blocks of transactions (e.g. minimum 1.000), with volume-based discounts.

5.2 Capacity Reservation model

Pay on a recurring basis to access Edge Computing Services (IaaS/PaaS/SaaS) using a preselected combination of service units (computing, storage, networking) that is permanently allocated to the Customer. Usually chosen for longer time periods in which the Customer has a permanent demand to attend. It is usually charged applying discounts to the pay-per-use tariffs that scale with the capacity volume that is allocated. Examples include:

- Monthly charges for certain functionality (e.g. compute or database)
- Per user charges
- Per API request volume

Note: It may include coverage options and fees.

5.3 Hybrid model

This includes a combination of reservation and pay per use where some resources are allocated for a certain period (capacity reservation model) and others on an on-demand basis (pay per use model).

5.4 Tiered Pricing models

Over the models described above, a tiered pricing can be applied to the unitary component consumed or reserved:

- QoS SLAs. For instance, based on:
  - guaranteed level of latency,
  - data transfer capacity and speed
  - data flow type (device to edge, edge to cloud, edge to edge)
  - additional services/features included
- Time of use (peak/off-peak hours)
- Based on computing and networking demand (e.g. uber-like)
- Based on markets/nodes/regions covered (following availability of alternative infrastructures)
- Based on connectivity or edge provider.
- Service based (e.g. security or AI frameworks)
- Network location (On-prem edge, far edge, near edge, central node)

5.5 Serverless model

In this model, the Customer requests to run a certain application with certain performance characteristics, and it is the system that decides the type and amount of resources that are required at each location to run it. The amount of resources scales up and down automatically with the demand of application usage, requiring neither the Customer nor the TEC Provider manual intervention.

This may be relevant for cases in which demand is uncertain and may vary significantly in volume and location over time (e.g. applications for users in mobility).

Charging and pricing for serverless model follow the same principles as described above for Pay-per-use and Tiered pricing models. Capacity reservation scheme does not apply in this case.

5.6 Spot/On-demand models

TEC offers a marketplace that allows their members to offer spot sales of unutilised capacity. TEC members would have visibility of this capacity and could bid for it using APIs (real-time dynamic market).

In this case, it follows a capacity reservation charging scheme, with pricing defined in the bidding process.

5.7 Other Considerations

Pricing may be adjusted based on some parameters:

- Volume discounts
- Multi-tenancy discounts
- Pre/post paid
- Flat rates for some specific applications that are not intensive in processing, storage or networking but require availability and prioritisation (e.g. public safety).

In some cases, services may be elected to be offered for free with procurement of other TEC services or for promotional purposes (free of charge temporary allocation of basic resources for demos).

6. Conclusion

This document describes the basic services the Telco Edge Cloud (TEC) will provide together with a set of commercial principles around stakeholder relationships and charging schemes. It intends to help Operators and service providers, technology providers and Customers understand the future TEC service portfolio and the value it will bring, assess the business potential in edge computing, and identify the way to contribute to and/or benefit from TEC.

As mentioned before, this whitepaper is complemented by the Operator Platform Telco Edge Proposal, which describes the technical framework and requirements of the platform that will support the Telco Edge Cloud service. The Operator Platform Group should refer to this whitepaper when describing further requirements in its future document versions, to be published in the coming months.

Further work is required to define the edge service in more detail and elaborate on the commercial principles. Commercial trials, several of which are already underway among operators and Customers, will provide valuable feedback in this direction. Engagement and collaboration with other industry associations will also contribute to build a good understanding of the Customer requirements. Additional knowledge gathered by operators in the TEC initiative may be added to a next iteration of this whitepaper and a potential future PRD.

References

Ref	Doc Number	Title
ETSI	GR NFV 003	Network Function Virtualisation; Terminology for Main Concepts in NFV
ETSI	GR NFV-IFA 029	Network Function Virtualisation Release 3; Architecture; Report on Enhancements of the NFV architecture towards “Cloud-native” and “PaaS”.
NIST	800-145	The NIST Definition of Cloud Computing
NIST	500-292	NIST Cloud Computing Reference Architecture
ITU-T	Y.3500	Cloud computing – Overview and vocabulary
ITU-T	Y.3501	Cloud Computing Framework & High-level Requirements
ITU-T	Y.3512	Cloud computing – Functional requirements of Network as a Service
ITU-T	Y.3515	Cloud computing – Functional architecture of Network as a Service
DMTF	DSP-ISO501	Software Defined Data Center (SDDC) Definition
GSMA	OPG.01	Operator Platform Telco Edge Proposal

Definitions and abbreviations

The following definitions and abbreviations have been used in this document.

Definitions

Term	Definition
Container Infrastructure Service Manager	CISM is a function that manages one or more container infrastructure services. These services provide runtime environment for one or more container virtualisation technologies. It can run on top of a bare metal or hypervisor-based virtualisation. Kubernetes is one of the most common CISMs.
Cloud Computing	Cloud computing is the on-demand availability of computer system resources, mainly storage and computing capabilities, without direct active management by the user. Cloud computing is delivered by a set of networked elements located in multiple data centers available to many users over the Internet. The elements are not individually addressed or managed by the users. Instead, the provider manages the entire suite of hardware and software that is presented to the user as an amorphous cloud, as a set of abstract or virtual computing capabilities.
Customer / Enterprise Customer	A company that uses the TEC to deploy edge applications in locations that are closer to the End Users than the traditional public cloud data centres.
Edge Node	In this document, it is referring to a point of presence in an Edge Cloud that hosts computing, storage and networking resources that are used to host Customers' applications and data, as defined in [OPG.01]. It is also named in other bodies as Cloudlet.
Edge Service Provider	A company providing commercial edge cloud services (potentially IaaS, PaaS, SaaS, and a marketplace) to market, aggregating multiple operator edge resources and exposing to its customers via its own platform.
End User	Person, or device, that runs an application that is connected to the Edge Cloud
Federation	In a generic way, a federation is a group of entities with a central government but independent in internal affairs. Federation is also the action of forming organisations into a single group with centralised control. Applied to edge computing, it represents the action of several operators and potentially service providers that cooperate so as to offer Customers a cohesive edge compute service across their combined footprint.
Hyperscaler	Cloud service provider that implements hyperscale computing, which is a distributed infrastructure that can scale to thousands of servers. The term is often associated with companies owning very large data centres like Microsoft, Facebook, Google, and Amazon. Some of these companies are starting to provide smaller data centres, for example at the edge of the network and on-premise..
Kubernetes	Kubernetes is a portable, extensible, open-source platform for managing containerised workloads and services that facilitates both declarative configuration and automation
Operator	A provider of fixed and/or wireless communications services that may include licence, ownership and operation of network infrastructure and associated systems and services, and radio spectrum allocation.

Operator Platform	Platform that facilitates access to the Edge Cloud capability of an Operator or federation of Operators by using a set of open and standard APIs. It follows the architectural and technical principles defined in the [OPG.01]. NOTE: Future versions of this document may extend the capabilities of the Operator Platform.
Service Providers	A company that provides solution development, system integration, IT, or technology operations services. Such companies include application providers, Operator enterprise business units and Hyperscalers.
Software Defined Networks	Network technology that enables dynamic, programmatically efficient network configuration in order to improve network performance and monitoring, using a cloud computing-like approach for network management. SDN applies a using a CUPS (Control-User Plane Separation) approach by disaggregating the forwarding functionality (data plane) from the routing functionality (control plane), that is hosted in a network component that centralises the network intelligence, the SDN controller. SDN presents a dynamic architecture that provides more flexibility and facilitates troubleshooting, but also may bring some challenges of scalability, security and elasticity.
Software Defined Storage	Software for policy-based provisioning and management of data storage independent of the underlying hardware. SDS implements storage virtualisation to separate the storage hardware from the software that manages it.
TEC Federation Hub	TEC component that implements the Federation Broker role defined by [OPG.01]
TEC Service Provider	Edge Service Provider that is part of TEC
Telco Edge Cloud	A type of the Operator Platform that exposes Edge Cloud capabilities (Naas, IaaS, and PaaS) to Customers and Service Providers.

Abbreviations

Term	Definition
AI	Artificial Intelligence
API	Application Programming Interface
CISM	Container Infrastructure Service Manager, as defined at ETSI GR NFV-IFA 029 or ETSI NG NFV 003
CPU	Central Processing Unit
DC-SDN	DataCenter Software Defined Networking
EWBI	East-Westbound Interface
FPGAs	Field Programmable Gate Arrays
GB	Gigabyte
GPU	Graphics Processing Unit
HA	High Availability
HDD	Hard Disk Drive
IaaS	Infrastructure as a Service, as defined in ITU-T Y.3500
IoT	Internet of Things

ISV	Independent Software Vendor
IT	Information Technology
KPI	Key Performance Indicator
MB	Megabyte
Mbps	Megabits per second
MEC	Multi-Access Edge Computing
ML	Machine Learning
NaaS	Network as a Service, as defined in ITU-T Y.3500
NPU	Network Processing Unit
OPG	Operator Platform Group
PaaS	Platform as a Service, as defined in ITU-T Y.3500
PRD	Permanent Reference Document
QoS	Quality of Service
RAM	Random Access Memory
SaaS	Software as a Service, as defined in ITU-T Y.3500
SBI	Southbound Interface
SDK	Software Development Kit
SDN	Software Defined Networks
SLA	Service-Level Agreement
SP	Service Provider
SSD	Solid State Drive
TEC	Telco Edge Cloud
VM	Virtual Machine

GSMA HEAD OFFICE
Floor 2
The Walbrook Building
25 Walbrook
London
EC4N 8AF
United Kingdom
www.gsma.com