

VERSION

2.0

STATE OF THE  
**EDGE**  
2019

*Open Glossary  
of Edge Computing*

**LF**EDGE

# *Open Glossary of Edge Computing, Version 2.0*

The [Open Glossary of Edge Computing](#) is an official project of [The Linux Foundation](#) and a founding project of [LF Edge](#). LF Edge is an umbrella organization within the Linux Foundation that aims to establish an open, interoperable framework for edge computing independent of hardware, silicon, cloud, or operating system.

For the past year, the [Open Glossary team](#) has been quietly updating the glossary in the [Github repo](#), incorporating feedback and addressing suggestions and concerns from within LF Edge, as well as the large community. The team has brokered relationships with other organizations, including the TIA and OpenStack Foundation, and now has working groups developing a taxonomy of edge computing and an edge computing landscape map.

As part of LF Edge, The Open Glossary leverages a diverse community to develop and improve upon this shared lexicon, offering an organization-and vendor-neutral platform for advancing a common understanding of edge computing and the next generation internet ecosystems. The project seeks to curate, define and harmonize terms related to the field of edge computing. Project participants submit common and accepted definitions into an openly licensed repository.

The Open Glossary is governed using a transparent and meritocratic process. Anybody can make additions, clarifications and suggestions by raising a GitHub issue or editing a branch and issuing a pull request. Each issue or pull request is evaluated by the community for inclusion.

The glossary is freely licensed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International license (CC-BY-SA-4.0), in order to encourage use and adoption. Code contributions to the project, such as scripts to build the crosslinks into the markdown file as well as scripts to produce a professional-looking PDF, are licensed under the Apache License, version 2.0 (Apache-2.0). These licenses are officially recorded in the project's LICENSE file.

- For information on how to contribute to the glossary: [See the Contributing Guide](#)
- To view a markdown version of the glossary: [See edge-glossary.md](#)
- To send a private email to the project Chair: [state-of-the-edge@gmail.com](mailto:state-of-the-edge@gmail.com)

## 3G, 4G, 5G

3rd, 4th, and 5th generation cellular technologies, respectively. In simple terms, 3G represents the introduction of the smartphone along with their mobile web browsers; 4G, the current generation cellular technology, delivers true broadband internet access to mobile devices; the coming 5G cellular technologies will deliver massive bandwidth and reduced latency to cellular systems, supporting a range of devices from smartphones to autonomous vehicles and large-scale IoT. Edge computing at the infrastructure edge is considered a key building block for 5G.

See also: *Infrastructure Edge*

## Access Edge Layer

The sub-layer of infrastructure edge closest to the end user or device, zero or one hops from the last mile network. For example, an edge data center deployed at a cellular network site. The Access Edge Layer functions as the front line of the infrastructure edge and may connect to an aggregation edge layer higher in the hierarchy.

See also: *Aggregation Edge Layer*

## Access Network

A network that connects subscribers and devices to their local service provider. It is contrasted with the core network which connects service providers to one another. The access network connects directly to the infrastructure edge.

See also: *Infrastructure Edge*

## Aggregation Edge Layer

The layer of infrastructure edge one hop away from the access edge layer. Can exist as either a medium-scale data center in a single location or may be formed from multiple interconnected micro data centers to form a hierarchical topology with the access edge to allow for greater collaboration, workload failover and scalability than access edge alone.

See also: *Access Layer Edge*

## Base Station

A network element in the RAN which is responsible for the transmission and reception of radio signals in one or more cells to or from user equipment. A base station can have an integrated antenna or may be connected to an antenna array by feeder cables. Uses specialized digital signal processing and network function hardware. In modern RAN architectures, the base station may be split into multiple functional blocks operating in software for flexibility, cost and performance.

See also: *Cloud RAN (C-RAN)*

## Baseband Unit (BBU)

A component of the Base Station which is responsible for baseband radio signal processing. Uses specialized hardware for digital signal processing. In a C-RAN architecture, the functions of the BBU may be operated in software as a VNF.

See also: *Cloud RAN (C-RAN)*

## Central Office (CO)

An aggregation point for telecommunications infrastructure within a defined geographical area where telephone companies historically located their switching equipment. Physically designed to house telecommunications infrastructure equipment but typically not suitable to house compute, data storage and network resources on the scale of an edge data center due to their inadequate flooring, as well as their heating, cooling, ventilation, fire suppression and power delivery systems. In the case when the hardware is specifically designed for edge cases it can cope with the physical constraints of Central Offices. See also: *Central Office Re-architected as Data Center (CORD)*

## Central Office Re-architected as Data Center (CORD)

An initiative to deploy data center-level compute and data storage capability within the CO. Although this is often logical topologically, CO facilities are typically not physically suited to house compute, data storage and network resources on the scale of an edge data center due to their inadequate flooring, as well as their heating, cooling, ventilation, fire suppression and power delivery systems. See also: *Central Office (CO)*

## Centralized Data Center

A large, often hyperscale physical structure and logical entity which houses large compute, data storage and network resources which are typically used by many tenants concurrently due to their scale. Located a significant geographical distance from the majority of their users and often used for cloud computing. See also: *Cloud Computing*

## Cloud Computing

A system to provide on-demand access to a shared pool of computing resources, including network, storage, and computation services. Typically utilizes a small number of large centralized data centers and regional data centers today. See also: *Centralized Data Center*

## Cloud Native Network Function (CNF)

A Virtualized Network Function (VNF) built and deployed using cloud native technologies. These technologies include containers, service meshes, microservices, immutable infrastructure and declarative APIs that allow deployment in public, private and hybrid cloud environments through loosely coupled and automated systems. See also: *Virtualized Network Function (VNF)*

## Cloud Node

A compute node, such as an individual server or other set of computing resources, operated as part of a cloud computing infrastructure. Typically resides within a centralized data center. See also: *Edge Node*

## Cloud RAN (C-RAN)

An evolution of the RAN that allows the functionality of the wireless base station to be split into two components: A Remote Radio Head (RRH) and a centralized BBU. Rather than requiring a BBU to be located with each cellular radio antenna, C-RAN allows the BBUs to operate at some distance from the tower, at an aggregation point, often referred to as a Distributed Antenna System (DAS) Hub (#distributed-antenna-system-das-hub). Co-locating multiple BBUs in an aggregation facility creates infrastructure efficiencies and allows for a more graceful evolution to Cloud RAN. In a C-RAN architecture, tasks performed by a legacy base station are often performed as VNFs operating on infrastructure edge micro data centers on general-purpose compute hardware. These tasks must be performed at high levels of performance and with as little latency as possible, requiring the use of infrastructure edge computing at the cellular network site to support them.

See also: *Infrastructure Edge, Distributed Antenna System (DAS) Hub*

## Cloud Service Provider (CSP)

An organization which operates typically large-scale cloud resources comprised of centralized and regional data centers. Most frequently used in the context of the public cloud. May also be referred to as a Cloud Service Operator (CSO).

See also: *Cloud Computing*

## Cloudlet

In academic circles, this term refers to a mobility-enhanced public or private cloud at the infrastructure edge, as popularized by Mahadev Satyanarayanan of Carnegie Mellon University. It is synonymous with the term Edge Cloud as defined in this glossary. It has also been used interchangeably with Edge Data Center and Edge Node in the literature. In a 3-tier computing architecture, the term “cloudlet” refers to the middle tier (Tier 2), with Tier 1 being the cloud and Tier 3 being a smartphone, wearable device, smart sensor or other such weight/size/energy-constrained entity. In the context of CDNs such as Akamai, cloudlet refers to the practice of deploying self-serviceable applications at CDN nodes.

See also: *Edge Cloud, Edge Data Center, Edge Node*

## Co-Location

The process of deploying compute, data storage and network infrastructure owned or operated by different parties in the same physical location, such as within the same physical structure. Distinct from Shared Infrastructure as co-location does not require infrastructure such as an edge data center to have multiple tenants or users.

See also: *Shared Infrastructure*

## Computational Offloading

An edge computing use case where tasks are offloaded from an edge device to the infrastructure edge for remote processing. Computational offloading seeks, for example, performance improvements and energy savings for mobile devices by offloading computation to the infrastructure edge with the goal of minimizing task execution latency and mobile device energy consumption. Computational offloading also enables new classes of mobile applications that would require computational power and storage



capacity that exceeds what the device alone is capable of employing (e.g., untethered Virtual Reality). In other cases, workloads may be offloaded from a centralized to an edge data center for performance. The term is also referred to as cloud offload and cyber foraging in the literature.

See also: *Traffic Offloading*

### Content Delivery Network (CDN)

A distributed system positioned throughout the network that positions popular content such as streaming video at locations closer to the user than are possible with a traditional centralized data center. Unlike a data center, a CDN node will typically contain data storage without dense compute resources. When using infrastructure edge computing CDN nodes operate in software at edge data centers.

See also: *Edge Data Center*

### Core Network

The layer of the service provider network which connects the access network and the devices connected to it to other network operators and service providers, such that data can be transmitted to and from the internet or to and from other networks. May be multiple hops away from infrastructure edge computing resources.

See also: *Access Network*

### Customer-Premises Equipment (CPE)

The local piece of equipment such as a cable network modem which allows the subscriber to a network service to connect to the access network of the service provider. Typically one hop away towards the end users from infrastructure edge computing resources.

See also: *Access Network*

### Data Center

A purpose-designed structure that is intended to house multiple high-performance compute and data storage nodes such that a large amount of compute, data storage and network resources are present at a single location. This often entails specialized rack and enclosure systems, purpose-built flooring, as well as suitable heating, cooling, ventilation, security, fire suppression and power delivery systems. May also refer to a compute and data storage node in some contexts. Varies in scale between a centralized data center, regional data center and edge data center.

See also: *Centralized Data Center*

### Data Gravity

The concept that data is not free to move over a network and that the cost and difficulty of doing so increases as both the volume of data and the distance between network endpoints grows, and that applications will gravitate to where their data is located. Observed with applications requiring large-scale data ingest.

See also: *Edge-Native Application*

## Data Ingest

The process of taking in a large amount of data for storage and subsequent processing. An example is an edge data center storing much footage for a video surveillance network which it must then process to identify persons of interest.

See also: *Edge-Native Application*

## Data Reduction

The process of using an intermediate point between the producer and the ultimate recipient of data to intelligently reduce the volume of data transmitted, without losing the meaning of the data. An example is a smart data de-duplication system.

See also: *Edge-Native Application*

## Data Sovereignty

The concept that data is subject to the laws and regulations of the country, state, industry it is in, or the applicable legal framework governing its use and movement.

See also: *Edge-Native Application*

## Decision Support

The use of intelligent analysis of raw data to produce a recommendation which is meaningful to a human operator. An example is processing masses of sensor data from IoT devices within the infrastructure edge to produce a single statement that is interpreted by and meaningful to a human operator or higher automated system.

See also: *Edge-Native Application*

## Device Edge

Edge computing capabilities on the device or user side of the last mile network. Often depends on a gateway or similar device in the field to collect and process data from devices. May also use limited spare compute and data storage capability from user devices such as smartphones, laptops and sensors to process edge computing workloads. Distinct from infrastructure edge as it uses device resources.

See also: *Infrastructure Edge*

## Device Edge Cloud

An extension of the edge cloud concept where certain workloads can be operated on resources available at the device edge. Typically does not provide cloud-like elastically-allocated resources, but may be optimal for zero-latency workloads.

See also: *Edge Cloud*

## Distributed Antenna System (DAS) Hub

A location which serves as an aggregation point for many pieces of radio communications equipment, typically in support of cellular networks. May contain or be directly attached to an edge data center deployed at the infrastructure edge.

See also: *Edge Data Center*

## Edge Cloud

Cloud-like capabilities located at the infrastructure edge, including from the user perspective access to elastically-allocated compute, data storage and network resources. Often operated as a seamless extension of a centralized public or private cloud, constructed from micro data centers deployed at the infrastructure edge. Sometimes referred to as distributed edge cloud.

See also: *Cloud Computing*

## Edge Computing

The delivery of computing capabilities to the logical extremes of a network in order to improve the performance, operating cost and reliability of applications and services. By shortening the distance between devices and the cloud resources that serve them, and also reducing network hops, edge computing mitigates the latency and bandwidth constraints of today's Internet, ushering in new classes of applications. In practical terms, this means distributing new resources and software stacks along the path between today's centralized data centers and the increasingly large number of devices in the field, concentrated, in particular, but not exclusively, in close proximity to the last mile network, on both the infrastructure and device sides.

See also: *Infrastructure Edge*

## Edge Data Center

A data center which is capable of being deployed as close as possible to the edge of the network, in comparison to traditional centralized data centers. Capable of performing the same functions as centralized data centers although at smaller scale individually. Because of the unique constraints created by highly-distributed physical locations, edge data centers often adopt autonomic operation, multi-tenancy, distributed and local resiliency and open standards. Edge refers to the location at which these data centers are typically deployed. Their scale can be defined as micro, ranging from 50 to 150 kW+ of capacity. Multiple edge data centers may interconnect to provide capacity enhancement, failure mitigation and workload migration within the local area, operating as a virtual data center.

See also: *Virtual Data Center*

## Edge Exchange

Pre-internet traffic exchange occurring at an infrastructure edge data center. This function will typically be performed in the edge meet me room of an infrastructure edge data center, and may operate in a supplemental or hierarchical fashion with traditional centralized internet exchange points if a destination location is not present at the edge exchange, as is the case with internet-bound traffic. An edge exchange may be used in an attempt to improve end-to-end application latency compared with a centralized internet exchange.

See also: *Internet Exchange Point (IXP)*



## Edge Meet Me Room

An area within an edge data center where tenants and telecommunications providers can interconnect with each other and other edge data centers in the same fashion as they would in a traditional meet me room environment, except at the edge.

See also: *Interconnection*

## Edge Network Fabric

The system of network interconnections, typically dark or lit fiber, providing connectivity between infrastructure edge data centers and potentially other local infrastructure in an area. These networks due to their scale and most frequent location of operation can be considered metropolitan area networks, spanning a distinct geographical area typically located in an urban center.

See also: *Edge Exchange*

## Edge Node

A compute node, such as an individual server or other set of computing resources, operated as part of an edge computing infrastructure. Typically resides within an edge data center operating at the infrastructure edge, and is therefore physically closer to its intended users than a cloud node in a centralized data center.

See also: *Cloud Node*

## Edge-Enhanced Application

An application which is capable of operating in a centralized data center, but which gains performance, typically in terms of latency, or functionality advantages when operated using edge computing. These applications may be adapted from existing applications which operate in a centralized data center, or may require no changes.

See also: *Edge-Native Application*

## Edge-Native Application

An application which is impractical or undesirable to operate in a centralized data center. This can be due to a range of factors from a requirement for low latency and the movement of large volumes of data, the local creation and consumption of data, regulatory constraints, and other factors. These applications are typically developed for and operate on the edge data centers at the infrastructure edge. May use the infrastructure edge to provide large-scale data ingest, data reduction, real-time decision support, or to solve data sovereignty issues.

See also: *Edge-Enhanced Application*

## Fog Computing

A distributed computing concept where compute and data storage resource, as well as applications and their data, are positioned in the most optimal place between the user and Cloud with the goal of improving performance and redundancy. Fog computing workloads may be run across the gradient of compute and data storage resource from Cloud to the infrastructure edge. The term fog computing was originally coined by Cisco. Can utilize centralized, regional and edge data centers.

See also: *Workload Orchestration*

## Gateway Device

A subcategory of the device edge, referring to devices on the device edge side of the last mile network which operate as gateways for other local devices, with the goal of aggregating and facilitating data transference between local devices, many of which are battery-operated and may operate for extended periods in a low-power state, and external entities such as a data analysis application operating inside an edge data center at the infrastructure edge.

See also: *Resource Constrained Device*

## Infrastructure Edge

Edge computing capability, typically in the form of one or more edge data centers, which is deployed on the operator side of the last mile network. Compute, data storage and network resources positioned at the infrastructure edge allow for cloud-like capabilities similar to those found in centralized data centers such as the elastic allocation of resources, but with lower latency and lower data transport costs due to a higher degree of locality to user than with a centralized or regional data center.

See also: *Device Edge*

## Local Breakout

The capability to put internet-bound traffic onto the internet at an edge network node, such as an edge data center, without requiring the traffic to take a longer path back to an aggregated and more centralized facility.

## Interconnection

The linkage, often via fiber optic cable, that connects one party's network to another, such as at an internet peering point, in a meet-me room or in a carrier hotel. The term may also refer to connectivity between two data centers or between tenants within a data center, such as at an edge meet me room.

See also: *Edge Meet Me Room*

## Internet Edge

A sub-layer within the infrastructure edge where the interconnection between the infrastructure edge and the internet occurs. Contains the edge meet me room and other equipment used to provide this high-performance level of interconnectivity.

See also: *Interconnection*

## Internet Exchange Point (IXP)

Places in which large network providers, among other entities, converge for the direct exchange of traffic. A typical service provider will access tier 1 global providers and their networks via IXPs, though they also serve as meet points for like networks. IXPs are sometimes referred to as Carrier Hotels because of the many different organizations available for traffic exchange and peering. The internet edge may often connect to an IXP.

See also: *Internet Edge*

## IP Aggregation

The use of compute, data storage and network resources at the infrastructure edge to separate and route network data received from the cellular network RAN at the earliest point possible. If IP aggregation is not used, this data may be required to take a longer path to a local CO or other aggregation point before it can be routed on to the internet or another network. Improves cellular network QoS for the user.

See also: *Quality of Service (QoS)*

## Jitter

The variation in network data transmission latency observed over a period of time. Measured in terms of milliseconds as a range from the lowest to highest observed latency values which are recorded over the measurement period. A key metric for real-time applications such as VoIP, autonomous driving and online gaming which assume little latency variation is present and are sensitive to changes in this metric.

See also: *Quality of Service (QoS)*

## Last Mile

The segment of a telecommunications network that connects the service provider to the customer. The type of connection and distance between the customer and the infrastructure determines the performance and services available to the customer. The last mile is part of the access network, and is also the network segment closest to the user that is within the control of the service provider. Examples of this include cabling from a DOCSIS headend site to a cable modem, or the wireless connection between a customer's mobile device and a cellular network site.

See also: *Access Network*

## Latency

In the context of network data transmission, the time taken by a unit of data (typically a frame or packet) to travel from its originating device to its intended destination. Measured in terms of milliseconds at single or repeated points in time between two or more endpoints. A key metric of optimizing the modern application user experience. Distinct from jitter which refers to the variation of latency over time. Sometimes expressed as Round Trip Time (RTT).

See also: *Quality of Service (QoS)*

## Latency Critical Application

An application that will fail to function or will function destructively if latency exceeds certain thresholds. Latency critical applications are typically responsible for real-time tasks such as supporting an autonomous vehicle or controlling a machine-to-machine process. Unlike Latency Sensitive Applications, exceeding latency requirements will often result in application failure.

See also: *Edge-Native Application*

## Latency Sensitive Application

An application in which reduced latency improves performance, but which can still function if latency is higher than desired. Unlike a Latency Critical Application, exceeding latency targets will typically not result in application failure, though may result in a diminished user experience. Examples include image processing and bulk data transfers.

See also: *Edge-Enhanced Application*

## Location Awareness

The use of RAN data and other available data sources to determine with a high level of accuracy where a user is and where they may be located in the near future, for the purposes of workload migration to ensure optimum application performance.

See also: *Location-Based Node Selection*

## Location-Based Node Selection

A method of selecting an optimal edge node on which to run a workload based on the node's physical location in relation to the device's physical location with the aim of improving application workload performance. A part of workload orchestration.

See also: *Workload Orchestration*

## Micro Modular Data Center (MMDc)

A data center which applies the modular data center concept at a smaller scale, typically from 50 to 150 kW in capacity. Takes a number of possible forms including a rackmount cabinet which may be deployed indoors or outdoors as required. Like larger modular data centers, micro modular data centers are capable of being combined with other data centers to increase available resource in an area.

See also: *Edge Data Center*

## Mobile Edge

A combination of infrastructure edge, device edge and network slicing capabilities which are tuned to support specific use cases, such as real-time autonomous vehicle control, autonomous vehicle pathfinding and in-car entertainment. Such applications often combine the need for high-bandwidth, low-latency and seamless reliability.

See also: *Infrastructure Edge*

## Mobile Network Operator (MNO)

The operator of a cellular network, who is typically responsible for the physical assets such as RAN equipment and network sites required for the network to be deployed and operate effectively. Distinct from MVNO as the MNO is responsible for physical network assets. May include those edge data centers deployed at the infrastructure edge positioned at or connected to their cell sites under these assets. Typically also a service provider providing access to other networks and the internet.

See also: *Mobile Virtual Network Operator (MVNO)*

## Mobile Virtual Network Operator (MVNO)

A service provider similar to an MNO with the distinction that the MVNO does not own or often operate their own cellular network infrastructure. Although they will not own an edge data center deployed at the infrastructure edge connected to a cell site they may be using, the MVNO may be a tenant within that edge data center.

See also: *Mobile Network Operator (MNO)*

## Modular Data Center (MDC)

A method of data center deployment which is designed for portability. High-performance compute, data storage and network capability is installed within a portable structure such as a shipping container which can then be transported to where it is required. These data centers can be combined with existing data centers or other modular data centers to increase the local resources available as required.

See also: *Micro Modular Data Center (MMDC)*

## Multi-access Edge Computing (MEC)

An open application framework sponsored by ETSI to support the development of services tightly coupled with the Radio Access Network (RAN). Formalized in 2014, MEC seeks to augment 4G and 5G wireless base stations with a standardized software platform, API and programming model for building and deploying applications at the edge of the wireless networks. MEC allows for the deployment of services such as radio-aware video optimization, which utilizes caching, buffering and real-time transcoding to reduce congestion of the cellular network and improve the user experience. Originally known as Mobile Edge Computing, the ETSI working group renamed itself to Multi-Access Edge Computing in 2016 in order to acknowledge their ambition to expand MEC beyond cellular to include other access technologies. Utilizes edge data centers deployed at the infrastructure edge.

See also: *Infrastructure Edge*

## Network Function Virtualization (NFV)

The migration of network functions from embedded services inside proprietary hardware appliances to software-based VNFs running on standard x86 and ARM servers using industry standard virtualization and cloud computing technologies. In many cases NFV processing and data storage will occur at the edge data centers that are connected directly to the local cellular site, within the infrastructure edge.

See also: *Virtualized Network Function (VNF)*

## Network Hop

A point at which the routing or switching of data in transit across a network occurs; a decision point, typically at an aggregating device such as a router, as to the next immediate destination of that data. Reducing the number of network hops between user and application is one of the primary performance goals of edge computing.

See also: *Edge Computing*

## Northbound vs Southbound (and east/west)

The direction in which data is transmitted when viewed in the context of a hierarchy where the cloud is at the top, the infrastructure edge is in the middle, and the device edge is at the bottom. Northbound and southbound data transmission is defined as flowing to and from the cloud or edge data center accordingly. Eastbound and westbound data transmission is defined as occurring between data centers at the same hierarchical layer, for purposes such as workload migration or data replication. This may occur between centralized or between edge data centers.

See also: *Virtual Data Center*

## Over-the-Top Service Provider (OTT)

An application or service provider who does not own or operate the underlying network, and in some cases data center, infrastructure required to deliver their application or service to users. Streaming video services and MVNOs are examples of OTT service providers that are very common today. Often data center tenants.

See also: *Mobile Virtual Network Operator (MVNO)*

## Point of Presence (PoP)

A point in their network infrastructure where a service provider allows connectivity to their network by users or partners. In the context of edge computing, in many cases a PoP will be within an edge meet me room if an IXP is not within the local area. The edge data center would connect to a PoP which then connects to an IXP.

See also: *Interconnection*

## Quality of Experience (QoE)

The advanced use of QoS principles to perform more detailed and nuanced measurements of application and network performance with the goal of further improving the user experience of the application and network. Also refers to systems which will proactively measure performance and adjust configuration or load balancing as required. Can therefore be considered a component of workload orchestration, operating as a high-fidelity data source for an intelligent orchestrator.

See also: *Workload Orchestration*



### Quality of Service (QoS)

A measure of how well the network and data center infrastructure is serving a particular application, often to a specific user. Throughput, latency and jitter are all key QoS measurement metrics which edge computing seeks to improve for many different types of application, from real-time to bulk data transfer use cases.

See also: *Edge Computing*

### Radio Access Network (RAN)

A wireless variant of the access network, typically referring to a cellular network such as 3G, 4G or 5G. The 5G RAN will be supported by compute, data storage and network resources at the infrastructure edge as it utilizes NFV and C-RAN.

See also: *Cloud RAN (C-RAN)*

### Regional Data Center

A data center positioned in scale between a centralized data center and an edge data center. Significantly physically further away from end users than an edge data center, but closer to them than a centralized data center. Also referred to as a metropolitan data center in some contexts. Part of traditional cloud computing.

See also: *Cloud Computing*

### Resource Constrained Device

A subcategory of the device edge, referring to devices on the device edge side of the last mile network which are often battery-powered and may operate for extended periods of time in a power-saving mode. These devices are typically connected locally to a gateway device, which in turn transmits and receives data generated by and directed to them from sources outside of the local network, such as a data analysis application operating in an edge data center at the infrastructure edge.

See also: *Gateway Device*

### Service Provider

An organization which provides customers with access to its network, typically with the goal of providing that customer access to the internet. A customer will usually connect to the access network of the service provider from their side of the last mile.

See also: *Access Network*

### Shared Infrastructure

The use of a single piece of compute, data storage and network resources by multiple parties, for example two organizations each using half of a single edge data center, unlike co-location where each party possesses their own infrastructure.

See also: *Co-Location*

## Software Edge

From a software development and application deployment perspective, the point physically closest to the end user where application workloads can be deployed. Depending on the application workload and the current availability of computing resources, this point may be at the device edge, but will typically be within the infrastructure edge due to its cloud-like capability to provide elastic resources.

See also: *Network Function Virtualization (NFV)*

## Throughput

In the context of network data transmission, the amount of data per second that is able to be transmitted between two or more endpoints. Measured in terms of bits per second typically at megabit or gigabit scales as required. Although a minimum level of throughput is often required for applications to function, after this latency typically becomes the application-limiting and user experience-damaging factor.

See also: *Quality of Service (QoS)*

## Traffic Offloading

The process of re-routing data that would normally be delivered inefficiently, such as over long distance, congested, or high cost networks, to an alternative, more local destination (e.g., a CDN cache) or on to a lower-cost or more efficient network. Local Breakout is an example of using edge computing for traffic offloading.

See also: *Local Breakout*

## Truck Roll

In the context of edge computing, the act of sending personnel to an edge computing location, such as to an edge data center, typically to resolve or troubleshoot a detected issue. Such locations are often remote and operate for the majority of the time remotely, without onsite personnel. This makes the cost other practical considerations of truck rolls a potential concern for edge computing operators.

## Vehicle 2 Infrastructure (V2I)

The collection of technologies used to allow a connected or autonomous vehicle to connect to its supporting infrastructure such as an machine vision and route finding application operating in an edge data center at the infrastructure edge. Typically uses newer cellular communications technologies such as 5G or Wi-Fi 6 as its access network.

See also: *Access Network*

## Virtual Data Center

A virtual entity constructed from multiple physical edge data centers such that they can be considered externally as one. Within the virtual data center, work-loads can be intelligently placed within specific edge data centers or availability zones as required based on load balancing, failover or operator preference. In such a configuration, edge data centers are interconnected by low-latency net-working and are designed to create a redundant and resilient edge computing infrastructure.

See also: *Edge Data Center*

### Virtualized Network Function (VNF)

A software-based network function operating on general-purpose compute re-sources which is used by NFV in place of dedicated physical equipment. In many cases, several VNFs will operate on an edge data center at the infrastructure edge.

See also: *Network Function Virtualization (NFV)*

### Workload Orchestration

An intelligent system which dynamically determines the optimal location, time and priority for application workloads to be processed on the range of compute, data storage and network resources from the centralized and regional data centers to the resources available at both the infrastructure edge and device edge. Workloads may be tagged with specific performance and cost requirements which determines where they are to be operated as resources that meet them are available for use.

See also: *Software Edge*

### xHaul ("crosshaul")

The high-speed interconnection of two or more pieces of network or data center infrastructure. Backhaul and fronthaul are examples of xhaul.

See also: *Interconnection*