ericsson.com/
distributed-cloud

Edge computing and 5G

Harnessing the distributed cloud for 5G success

June 2019

Harnessing the edge opportunity is key for 5G success

With increasing interest in new use cases like smart manufacturing and augmented/virtual reality (AR/VR), as well as upcoming 5G radio networks, there is a clear need for distributed networks.

By 2023, 5G will make up around one-fifth of all mobile data traffic, with subscription uptake forecast to reach 1 billion. 4G LTE has also exploded since its launch in 2010, accounting for some 50 percent of total traffic in 2017 and still growing. Distributed cloud computing is paving the way for the future of network communications and, rather than waiting for 5G, operators should build tomorrow's networks today.

What is the edge?
Edge computing provides execution resources (compute and storage) with sufficient connectivity (networking) at close proximity to the data sources, typically within or at the boundary of access networks. Additionally, deployments within the customer premises managed or hosted by the operator are possible, for example in homes, enterprises and factory floors, as well as vehicles, including trains, planes and private cars. The core benefits of edge solutions are low latency, high bandwidth, and trusted computing and storage.

How about the public clouds?
Public cloud service providers will have to make major investments to build their own edge infrastructure. This presents an opportunity for operators to play a role in delivering these new services by providing the following assets:

- Geographical presence and distribution of sites at cell towers as well as sites across the network, creating sweet spots for every deployment
- Value-added connectivity with QoS capabilities – not just pipes but smart connectivity with the ability to steer, secure and break out traffic
- NFV-enabled network, providing the ability to place initial workloads and grow into the edge

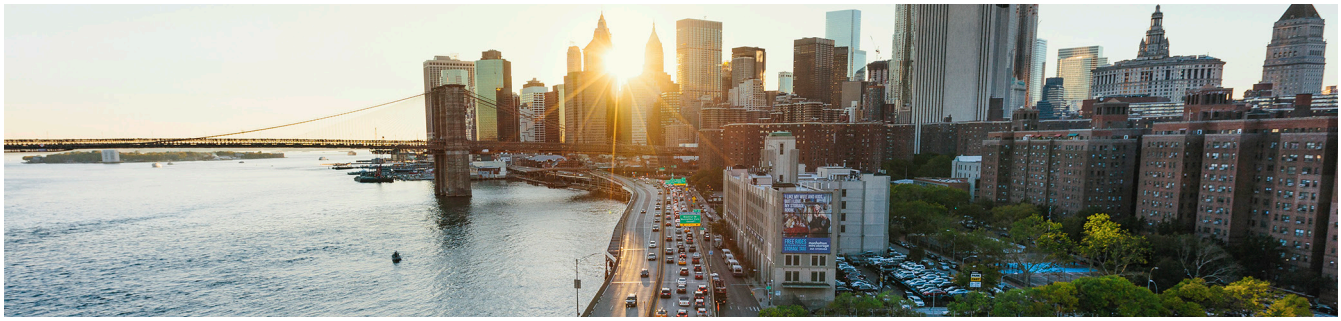
Solutions complementing public clouds are necessitated by use cases with high requirements on latency, bandwidth and local survivability/reliability. Some providers use platforms like Greengrass (AWS) or IoT Edge (Azure) that can be deployed at distributed sites and integrated seamlessly into the public cloud infrastructure and APIs. Others are rolling out highly distributed infrastructures; for example, Google has deployed more than 1,000 Google Global Caches. However, none of these solutions supports the vast number of potential mobile use cases where the computing capacity would have to be deployed between the antennas and the current peering points. These use cases require two things:

- Local breakout of the traffic close to the user and to the compute – this is not available today, but is something that only operators can provide
- A compute infrastructure to run the platforms – this can either be operator-owned or public cloud-owned infrastructure

Additionally, integrating with hundreds of operators will be cumbersome for public clouds, commercially and technically.

5G use cases and revenue potential for edge computing

Ericsson conducted an extensive study on the 5G business potential for operators showing that, between 2020 and 2024, the 5G revenue potential for operators will grow from USD 14 to 129 billion.[1]



Distributed cloud is suitable for use cases with demanding requirements for scalability, security, availability, low latency and bandwidth. These use cases account for 25 percent of the 5G business potential, and therefore revenue potential.

Capturing edge business opportunities
When addressing edge opportunities, one of the key aspects is how to work with the application ecosystem to provide solutions for industries. There are three ways for communications service providers (CSPs) to do this:

- Work with industries to drive the ecosystem to provide solutions
- Drive the ecosystem on their own
- Leverage the hyperscale cloud providers' ecosystems

Which model is best depends on the strategy and market position of the CSP, the type of industry being addressed, and how far their ecosystems have developed already.

The assets held by CSPs make them uniquely placed to contribute with essential capabilities based on distributed cloud and exposure of cloud, network and orchestration logic. Since many developers don't know how telecom networks work, it is important to simplify API exposure as much as possible. CSPs should aim for a converged solution with cloud execution, service exposure, and dynamic workload placements through industry-aligned APIs to best meet the needs of both industries and developers.

To succeed, operators also need to attract content and application developers. This is possible in several ways. Developers could be approached either directly or through a facilitator such as Ericsson's Edge Gravity. Another possibility is direct collaboration with public cloud providers which have existing relationships with developers, but no distribution capabilities.

Other benefits we see are transport and processing offload, which will be significant with the increased traffic of a 5G world, as well as increased efficiency of connected devices that are offloading to the edge.

The business benefits are already apparent: we have systems in commercial production with several tier 1 and tier 2 customers. These customers experienced dramatic improvements in the overall time-to-market of their network services, with cost reductions of up to 75 percent.

Demand driver	Edge capability in 5G
Application latency	With the app closer to the user and 5G radio, the latency can be <1ms and benefit new use cases.
Application exposure	The new 5G core will also offer app exposure for edge deployments.
Transport offload	5G bandwidths may increase traffic further. Service delivery from the edge will minimize the backhaul traffic.
Processing offload	Application processing at the edge will offload devices at central datacenters while preserving user experience.

25%
Edge computing will be vital for 25 percent of use cases.

[1] Ericsson, The guide to capturing the 5G-IoT business potential (2018)

Challenges and requirements for implementing edge solutions

There is some uncertainty about how to successfully overcome the challenges of implementing edge solutions.

Where to best place the applications?
From a cost perspective, centralization is generally preferred. However, the location depends on the use case and the value it provides, therefore the edge cannot be a fixed point.

In the example of AR, pictures from a phone are sent to a datacenter. This activity is neither latency-sensitive nor bandwidth-intense, so a centralized datacenter is best. Then augmentation data like videos and 3D models are sent to the device, a bandwidth-demanding task requiring content delivery network (CDN) functionality, which is best placed at distributed sites. Finally, videos or pictures can be augmented in the network, which is very latency-sensitive and requires close proximity.

Data-heavy applications require broader distribution to optimize bandwidth consumption. Key requirements are reliability and availability for applications that must always work, even if connectivity to central sites fails.

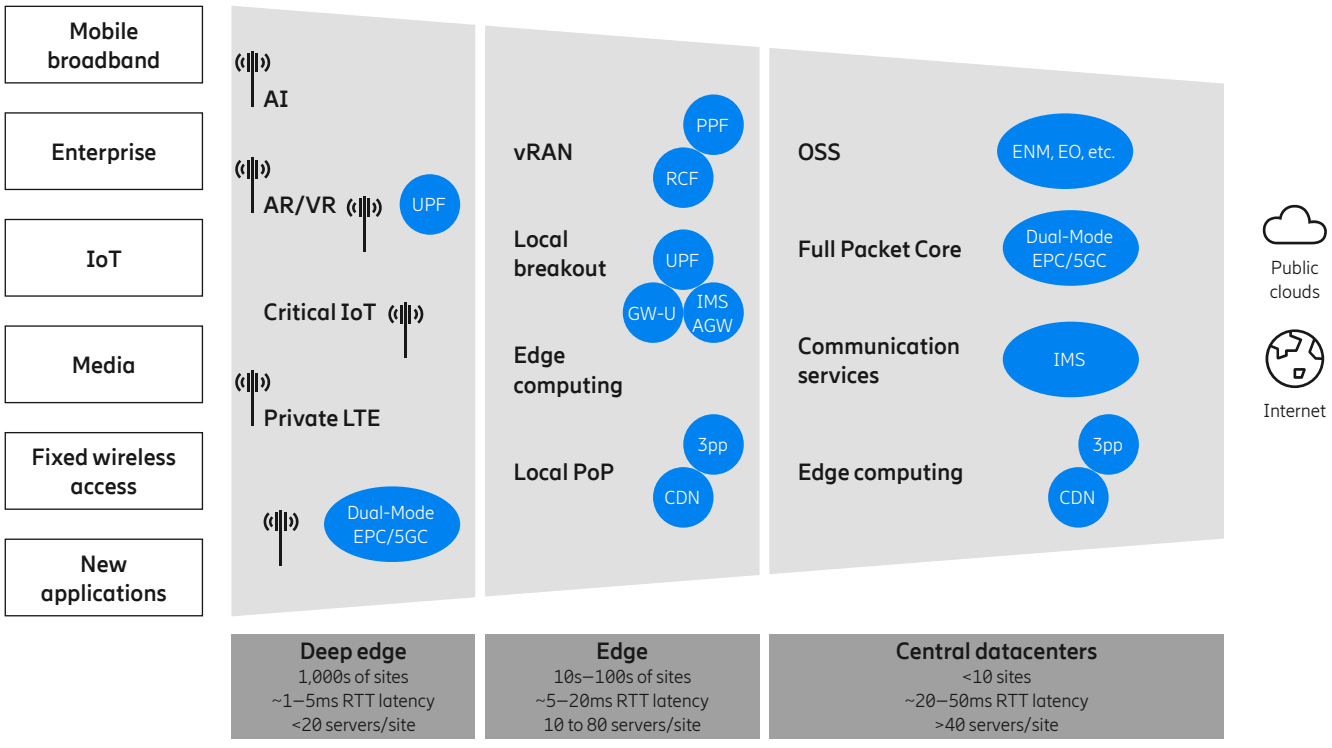
The function for location of the applications should be integrated in the overall service and network orchestration to be used for NFV applications, such as 5G Core and virtualized RAN functions (vRAN). This is important to facilitate automatic placement and network slice orchestration.

Exposure and application development
There must be a solution to enable application development and onboarding. The developers need not know the details of the distributed cloud and telecom application program interfaces (APIs), or the geographical location of edge locations in the network, in order to secure the SLAs.

Discussions with customers and application providers have confirmed that using computing at the edge must be as simple as using public cloud services.

Cloud native applications
5G and IoT unleash a variety of workloads with different characteristics and needs. Platforms must therefore be able to cater for the use of different deployment models supporting both virtual machine and container-based workloads.

Service providers must distribute to meet the needs of 5G applications



Good business models and API exposure for global reach
Historically, application developers were required to learn telecom network architectures and standards, instead of simplifying the representation of network assets. The challenges of bringing solutions without an adequate business model for the exposure (such as operators taking only individual approaches and a lack of business APIs) have caused service exposure attempts to fail during the last 20 years.

Operational efficiency and support
In NFV, the number of re-certification/validation cycles required dramatically increases due to the dynamic nature of the overall software stack, including NFVI, multi-vendor VNFs and support functions. Without automation, test organizations will become overwhelmed, causing reduced agility and potentially lower quality. A related challenge is how to locate faults and root causes in the entire stack of NFVI and VNFs.

The edge solution must have the capability to locate and handle faults quickly, in the entire stack and across a multi-vendor and multi-tenant environment. In cloud-based systems, isolating an issue to a particular VNF is no longer sufficient. The performance of each VNF is highly dependent on the underlying infrastructure layers, management layers, and the overlay network which may comprise of multiple vendors. It should be possible for the same operational model and team to manage both the edge infrastructure and the central NFV solution.

Management, orchestration and LCM
With potentially hundreds of edge locations like central offices, radio base stations or on enterprise premises, management will be a challenge if these systems are managed locally and all the layers of the stack are managed separately. Also, as we move to NFV we are looking at having CI/CD loops for cost and efficiency reasons.

The possibility to remotely monitor performance and perform low-touch life-cycle management activities such as software installations, restarts and upgrades is crucial to avoid expensive site visits. A software-defined infrastructure (SDI) capability can greatly improve the efficiency and utilization of the compute and storage resources.[2] This should have the ability to discover and manage all infrastructure elements such as configuration and control.

Data and regulatory compliance
With this new business need for flexible cloud solutions, data handling is just as important as computing, and regulatory compliance is an urgent driver. Operators may have to navigate regulations restricting the movement of data between geographical locations such as a country, a region or enterprise premises.

Infrastructure deployment cost and TCO
Most current NFV infrastructures are built for larger datacenter deployments. The reduced floor space and cooling capacity in locations closer to the network edge place high demands on hardware efficiency and footprint.

It will be crucial for the success of edge deployments to maintain a competitive total cost of ownership (TCO) for edge sites. We need to support systems of different sizes with minimal hardware and/or storage units. We also need to be able to scale efficiently depending on business needs and reuse existing site and networking assets.

The servers need to have a small footprint, low power requirements and be NEBS3 compliant to cater for extreme environmental conditions without cooling.

Scaling
Whilst most NFV-type deployments today consist of a few datacenters set up for redundancy and basic geographical distribution, the amount of sites will dramatically increase with the introduction of edge infrastructure. It is essential that the infrastructure management solution, the networking concept and the orchestration are capable of handling the scale of managed sites.

Solution integration can be lengthy and costly
Investments in hardware and software are significant, the main costs likely being the integration and deployment of the solution. The early NFV adopters in the telco industry embraced a best-of-breed approach that turned out to be far more time consuming and costly than first anticipated.

[2] Ericsson, SDI - software defined infrastructure (2018)

The edge infrastructure solution for a distributed telco cloud

Ericsson has introduced the distributed cloud to support all workloads as an extension of the Ericsson NFVI for telco cloud.

Overall solutions for edge computing and distributed cloud
The distributed cloud provides an execution environment for cloud application optimization across multiple sites, managed as one solution and perceived as such by the applications. It is based on SDN, NFV and 3GPP edge computing with end-to-end service orchestration and external exposure and management interfaces. This enables multi-access and multi-cloud capabilities and unlocks networks as open platforms for application innovations.

Ericsson NFVI includes Ericsson Cloud SDN built on OpenDaylight SDI, including hardware for central, distributed and edge sites (based on Open Compute Platform) and Cloud Execution Environment based on OpenStack and Ericsson Orchestrator (cloud orchestrator NFVO).[3]

Efficient deployment with Ericsson Edge NFVI
Our Edge NFVI solution is optimized for edge deployments. A 30–50 percent reduction in overheads reduces hardware investments through the need for fewer servers. Further optimization can also be achieved through Ericsson's Edge server with a shorter form factor. This form factor allows for installation in existing central offices as well as Ericsson radio sites in the same cabinets as the Ericsson baseband. Our Edge NFVI has the following key features:

- Small and powerful hardware that is NEBS3 compliant with integrated gateway router
- Optimized for dual-mode 5G core applications
- Fully orchestrated VNFs and life-cycle management, as well as SDI of the hardware platform

Low TCO solution integration
As we have learned from several complex and challenging NFVI integrations, a better approach than trying to integrate your own optimized components is to rely on a pre-configured, pre-integrated and tested, system-verified solution from an experienced partner.

While pre-integrated, Ericsson's Edge solution retains the flexibility to choose hardware and is open for any type of VNF through our partner certification program.

A vast number of parameters, in the range of thousands, are pre-set to achieve maximum performance. We use automated scripts to configure the solution to the customer's requirements. This has resulted in dramatic improvements in the time-to-market of network services, with cost reductions up to 75 percent.

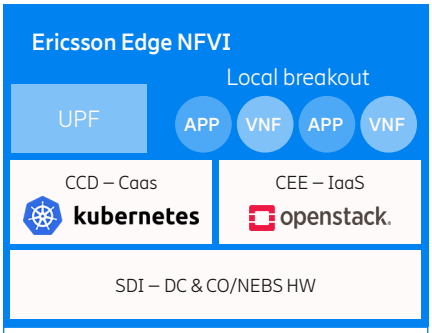
Exposure and business models
Ericsson's exposure approach is based on two main actions:

- Simplifying the exposure of network assets and capabilities (network slices, mobile network connectivity and anchoring points and infrastructure resources) by an approach that is closer to users and by hiding complexity, e.g. exposing a simple parametrization of network slices and adopting known standards in the industry like Kubernetes and ONAP.
- Working closely with developers and industry initiatives on API definition and use cases for ease of adoption, e.g. within AECC. Developers don't have to learn new telecom-specific APIs.

Edge Gravity
Ericsson Edge Gravity is a live global edge computing network in partnership with close to 100 service providers. Gravity started in 2016 with video streaming and content acceleration. The solution is designed to help service providers create a global network that is available to content providers and application providers who need more capacity and better performance for their consumer applications.[4]

Open source
Open source is starting to be seen by some to be the standard for composable cloud infrastructure.

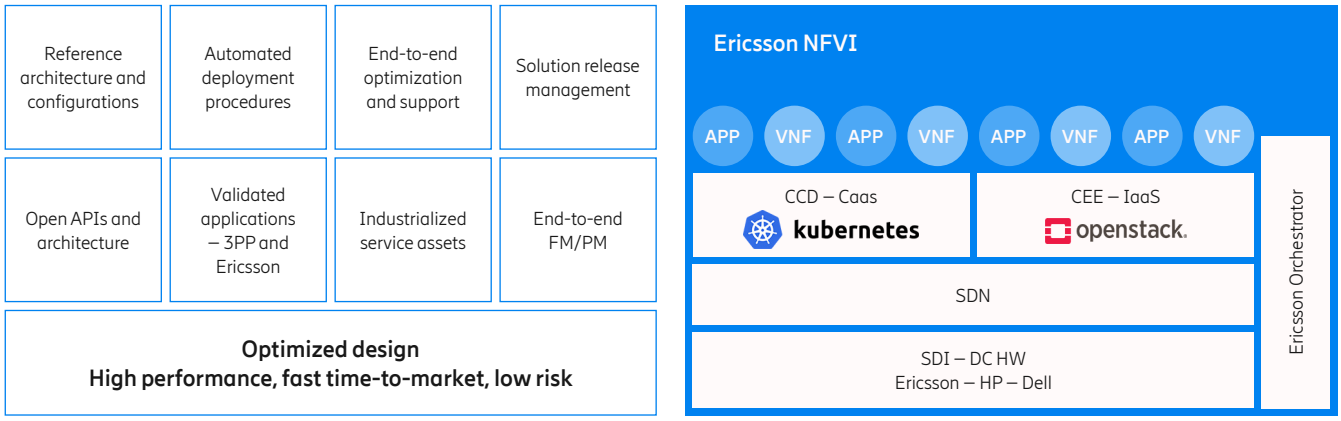
Open source communities like CNCF and projects like Kubernetes provide key ingredients of an edge software stack.



- Dual-Mode 5G Cloud Core is needed to connect local applications
- Optimized for distributed 4G and 5G with high throughput requirements
- Local breakout for selected devices, services or IP flows
- Run VNFs and edge applications on the same platform
- Manage VNFs, applications, and infrastructure from a single centralized pane of glass

[3] Ericsson, Build cloud & NFV infrastructure (2019)
[4] Ericsson, Edge gravity by Ericsson - Edge cloud platform (2019)

What we do to make an NFVI system-verified solution



Edge solution management and orchestration

The Ericsson NFVI and Edge NFVI solutions can be enhanced with the flexible Dynamic Orchestration solution to be able to manage the distribution of network functions, such as packet core or communication services, and the orchestration of network slicing in 5G. Orchestration is vital for distributed cloud, providing end-to-end management of networking, cloud infrastructure, workloads and network services. The solution is ONAP compatible.

Our edge solutions enable management of a large number of sites through an end-to-end dashboard. Automated workflows for common operational tasks such as upgrades, restarts or addition /removal of devices will be developed with a one-click approach in mind.

VNFs are managed by the Ericsson Cloud Deployment Engine (ECDE) platform, which is used for streamlining and automating NFV delivery and operations.

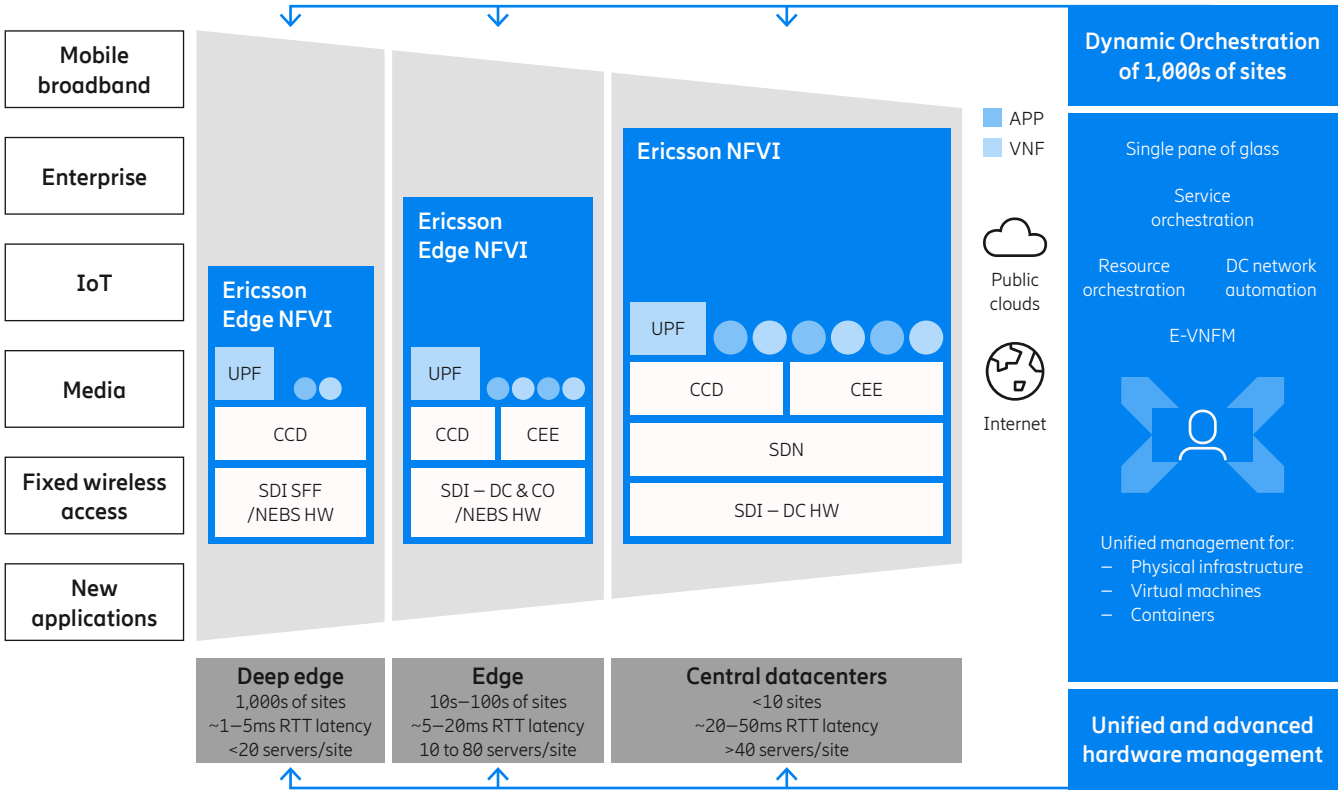
ECDE is designed to automate VNF onboarding and validation workflows. It provides capabilities to simplify setup as well as to track, operate and manage the VNF onboarding and validation processes using self-serve web portals and management dashboards.

Ericsson also supports operators through their network and operational transformation to telco cloud.

End-to-end and multi-vendor solution support

As well as helping to remedy and restore Ericsson products that may be experiencing trouble, we can help to isolate where the issue is occurring, whether it is an Ericsson issue or not. Our AI-based tools, combined with experienced engineers, will trace and restore issues across multiple vendors.

We provide end-to-end edge computing with distributed cloud



Ericsson enables communications service providers to capture the full value of connectivity. The company's portfolio spans Networks, Digital Services, Managed Services, and Emerging Business and is designed to help our customers go digital, increase efficiency and find new revenue streams. Ericsson's investments in innovation have delivered the benefits of telephony and mobile broadband to billions of people around the world. The Ericsson stock is listed on Nasdaq Stockholm and on Nasdaq New York.

www.ericsson.com

Ericsson
SE-164 80 Stockholm, Sweden
Telephone +46 10 719 0000
www.ericsson.com

EAB-19:004157 Uen