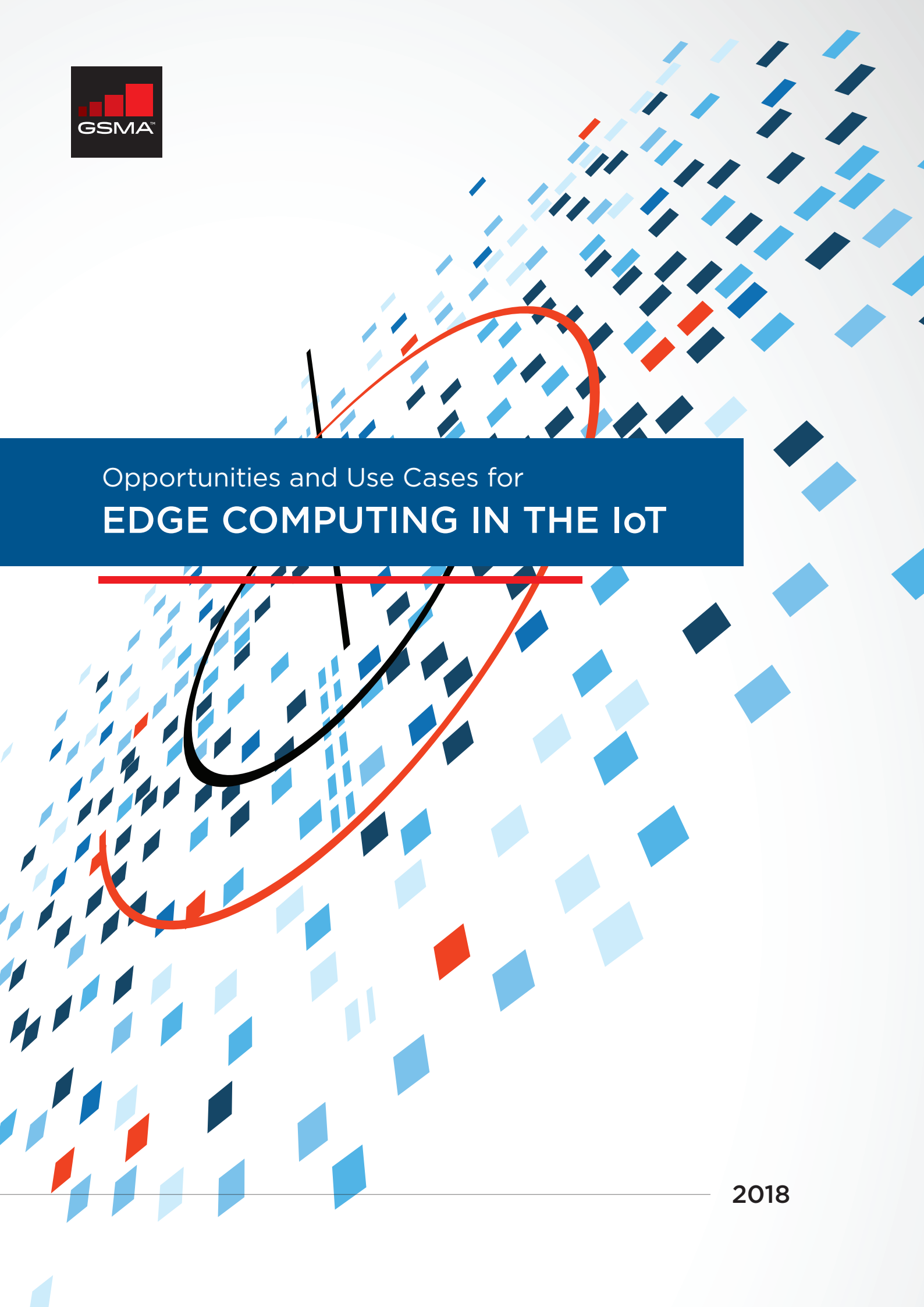


Opportunities and Use Cases for

EDGE COMPUTING IN THE IoT

2018

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 GSMA Internet of Things Programme

The GSMA's Internet of Things Programme is an industry initiative focused on:

- COVERAGE of machine friendly, cost effective networks to deliver global and universal benefits
- CAPABILITY to capture higher value services beyond connectivity, at scale
- CYBERSECURITY to enable a trusted IoT where security is embedded from the beginning, at every stage of the IoT value chain developing key enablers, facilitating industry collaboration and supporting network optimisation, the Internet of Things Programme is enabling consumers and businesses to harness a host of rich new services, connected by intelligent and secure mobile networks.

Visit gsma.com/iot to find out more.

1. Executive Summary

Edge computing for the Internet of Things (IoT) allows IoT deployments to be enhanced through data processing closer to the end device. This results in lower latency and improved efficiencies in data transport.

IoT edge computing is significantly different from non-IoT edge computing, with distinct demands and considerations. IoT devices typically have limited data processing and storage capabilities, so substantial data processing needs to occur off the device, with the edge offering an environment to undertake this processing and manage large volumes of IoT devices and data. This in turn can reduce device cost, as many functions can be off-loaded to the edge. The location of the edge itself has various possibilities and will differ according to the use case. For example, the edge for IoT could reside at an operator's local or regional data centre, at a base station or at a dedicated server on the customer's premises.

The use cases for IoT edge that have been prioritised for exploration by the GSMA are shown in the table below:

Category	IoT FOUNDATION Supporting general IoT functionality	IoT SERVICE ENABLERS Adding value to IoT	IoT SOLUTIONS Supporting customer solutions
Use Case	◢ Device Management ◢ Security	◢ Priority Messaging ◢ Data Aggregation ◢ Data Replication ◢ Cloud Enablement	◢ IoT Image and Audio Processing

TABLE OF CONTENTS

2. Introduction

IoT market analysts expect the edge to play a significant role in supporting IoT implementations going forward, as it creates efficiencies and scale in networks that makes IoT deployments more self-sustaining. IDC (International Data Corporation) estimate that that by 2020, IT spending on edge infrastructure will reach up to 18 per cent of the total spend on IoT infrastructure.[1] Mobile operators have the demonstrable capability to manage infrastructure, data and applications for IoT services, and are well placed to continue this with edge for IoT.

Rapidly increasing numbers of IoT devices and resultant data, mean that new techniques to meet customer requirements and ensure effective management need to be explored. Alternatives to the traditional IoT model of sending all data to the cloud for processing are required as the volume of data to be processed explodes, and the cost of centrally storing and processing every piece of data, important or not, becomes harder to justify. This is particularly important for IoT services, as they can generate large volumes of new data for analysis. Edge computing is a deployment model which aims to push the relevant data processing and storage attributes closer to where the device is located. This means that data can be processed more efficiently, and many attributes do not need to be centralised. Mobile operators are well placed to enable edge computing to scale and enhance IoT deployments, additionally allowing options for data processing on behalf of customers to be further incorporated into their service offerings.

However, implementation of edge computing for IoT by mobile operators is not without challenges. New investments, infrastructure and management platforms may be needed. Today, edge computing is a relatively immature technology that has been dominated by traditional cloud providers. Other members of the IoT ecosystem, including mobile operators, do have roles to play but as the IoT is still evolving, it will take some time for appropriate commercial and technical models to emerge.

This paper explores the benefits to the IoT of edge computing and some of the different use cases where it could be applied. It explores the potential operator role for IoT edge computing and identifies some potential next steps to be undertaken by the industry.

[1] IDC FutureScape: Worldwide IoT Predictions 2018

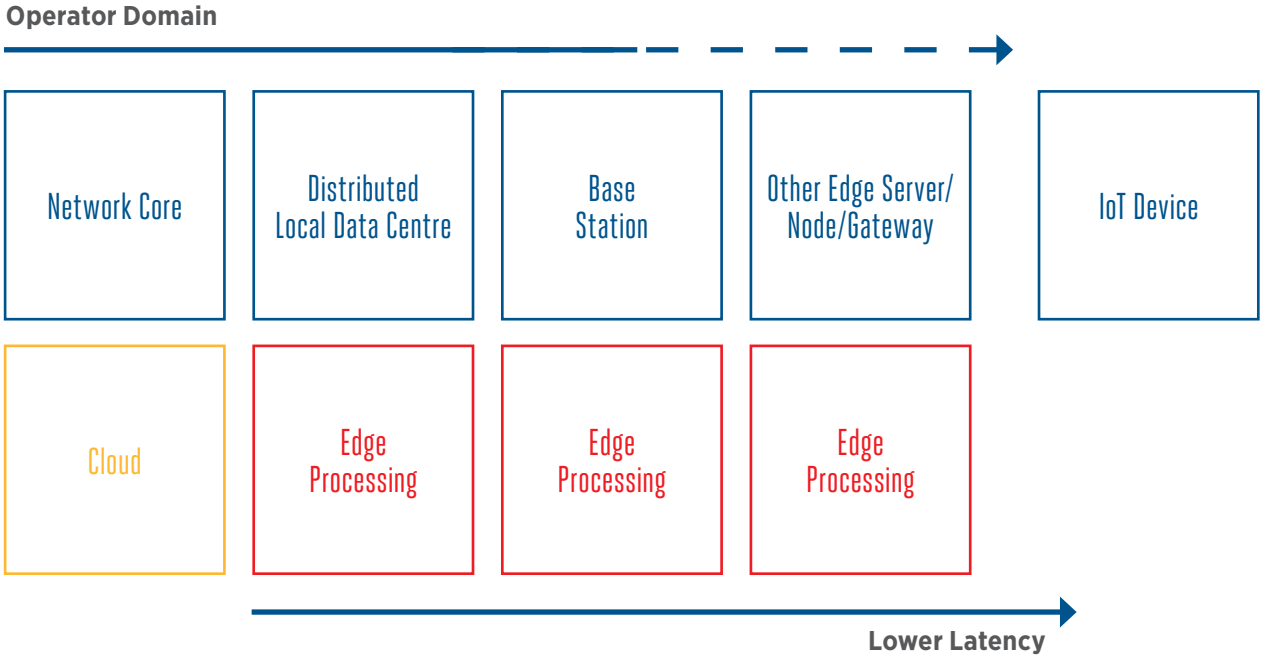
3. Where is the Mobile Operator IoT Edge?

The position of the edge for IoT can change depending on the service required by the customer. Services requiring very low latency will need a different edge than those with less urgent data processing requirements.

Mobile operators should view IoT edge computing as a flexible, distributed processing point beyond the core where network control, application logic, device management and data processing and analytics can be separated to enable a wide variety of deployment models along with automated and efficient management of IoT devices and data.

LOCATION OF THE IOT EDGE FROM A MOBILE OPERATOR POINT OF VIEW

THE MOBILE OPERATOR EDGE



4. Key Benefits of Edge for the IoT

Edge for the IoT brings potential benefits for many IoT deployments, including decreased response time along with increased communications efficiency, compared to using the cloud to process and store data. For example, many IoT processes can have a high level of automation at the edge resulting in low latency for rapid data processing. Only the most important information need then be sent to the cloud for further action or investigation.

Many new IoT services, such as intelligent vehicles, drones or smart grids, could come to rely on edge computing. Many of the benefits of IoT edge will need to be refined in proof of concept deployments by mobile operators to demonstrate that the model is beneficial. Benefits of IoT edge computing that have been identified include:

- **Low latency.**

By its nature, the edge is closer to the IoT device than the core or cloud. This means a shorter round-trip for communications to reach local processing power, significantly speeding up data communications and processing.

- **Longer battery life for IoT devices.**

Being able to open communication channels for shorter periods of time due to improved latency, means that battery life of battery powered IoT devices could be extended. distributed ledger, or a hybrid open source ledger implementation such as BigchainDB could be used to obtain the advantage of a distributed ledger which provides features from the NoSQL database MongoDB on which it is based.

- **More efficient data management.**

Processing data at the edge makes simple data quality management such as filtering and prioritisation more efficient. Completing this data administration at the edge, means cleaner data sets can be presented to cloud based processing for further analytics.

- **Access to data analytics and AI.**

Edge processing power and data storage could all be combined to enable analytics and AI, which require very fast response times or involve the processing of large ‘real-time’ data sets that are impractical to send to centralised systems.

- **Resilience.**

The edge offers more possible communication paths than a centralised model. This distribution means that resilience of data communications is more assured. If there is a failure at the edge, other resources are available to provide continuous operation.

- **Scalability.**

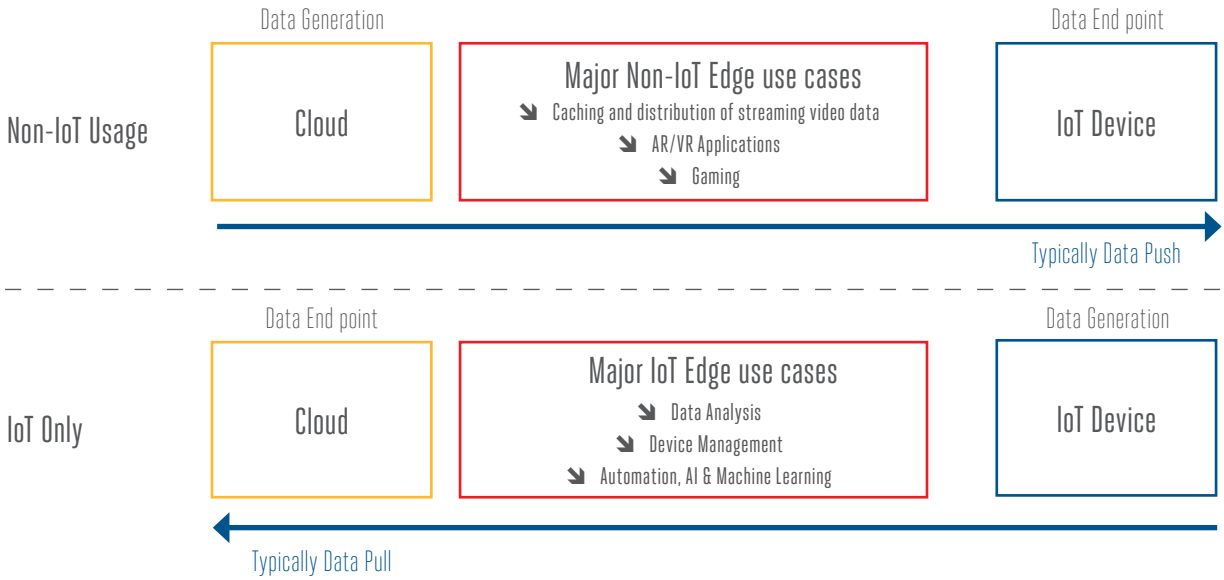
As processing is decentralised with the edge model, less load should ultimately be placed on the network. This means that scaling IoT devices should have less resource impact on the network, especially if application and control planes are located at the edge alongside the data.

5. Unique Requirements of Edge for the IoT

IoT edge requirements are different to those from non-IoT edge computing use cases. The IoT needs to support a large number of devices, many of which do not have their own dedicated data processing resources, but that may be generating a large volume of data.

The relationship between these devices and the edge is different from that of other connected devices such as smartphones, where much of the data processing can be completed on the device. In non-IoT use cases, the edge is often used to serve constant volumes of data to the end device enable services such as video streaming, or offer low latency applications for VR and gaming.

DIFFERENCES BETWEEN EDGE FOR IoT AND NON-IoT



For some customers, the IoT and other services will possibly share the same physical infrastructure and enablement platforms as non-IoT services. This means that considerations for the IoT need to be taken in the deployment of any edge infrastructure. If the IoT is to share edge resources with non-IoT services, then it may be necessary to have dedicated resources allocated on the edge node or gateway to support required levels of service. IoT services often run 24/7 while other edge use cases may only run intensively at peak times of day. Therefore, there needs to be consideration of the scope of IoT requirements on resources to ensure they can be met, even at peak times.

6. Use Cases for IoT Edge

The following use cases have been identified as some of those which may be able to benefit from IoT edge computing to improve the level of service for IoT devices and applications.

Category	IoT FOUNDATION Supporting general IoT functionality	IoT SERVICE ENABLERS Adding value to IoT	IoT SOLUTIONS Supporting customer solutions
Use Case	◢ Device Management ◢ Security	◢ Priority Messaging ◢ Data Aggregation ◢ Data Replication ◢ Cloud Enablement	◢ IoT Image and Audio Processing

6.1 IoT FOUNDATION

6.1.1 Device Management

There are many device attributes and configurations that can be controlled at the edge, with many device management platforms extending their functionality to manage devices connected to edge infrastructure. Below are examples of four different attributes that device management at the edge will need to support:

- Distributed firmware updates – the use of the edge gateway to distribute firmware updates locally, with distribution being managed by the edge node as opposed to the queuing system typically used when a firmware update is distributed centrally.
- Device configuration updates – devices at the edge will need to be configured locally as services change. The edge could be able to manage this remotely.
- Diagnostics of connected devices – the use of analytics at the edge can be used to identify specific problems with devices in the field through machine learning or pattern recognition.
- Edge node or gateway management – device management platforms can be used to manage the operator's edge infrastructure as well as the IoT device.

Benefits of the Edge to this Use Case

- **Better control at the edge**

Device management at the edge means that IoT devices are better controlled resulting in increased efficiency and improved quality of service. Where constraints are identified in edge computing, additional resources can be brought online to reduce problematic areas and better manage service scalability. Additionally, IoT services such as data access by third parties can be managed effectively if control is available at the edge node.

- **Better management of device performance**

Enforced consistency of device configurations and performance criteria at the edge mean that there are fewer variables to contend with, which means in turn that applications can be confidently optimised to obtain best the best performance. Additionally, device performance data can be collected effectively at the edge for further analysis.

- **Effective deployment of applications at the edge**

The use of a device management platform at the edge means that applications can be distributed to appropriate edge locations. If an IoT deployment has a number of different edges to manage, an application can be deployed to the edge closest to the device to obtain lowest latency, if a customer is paying for a premium service.

- **More efficient distribution of device updates**

Distributing patches and firmware updates to devices is a key feature of a device management platform. Bringing edge infrastructure into the ecosystem means that these updates can be made

more efficient as the edge can better manage local network resources for distribution to local devices.

Integrated hardware and software ecosystem

Device management that manages all aspects of the edge, including devices, applications and connectivity means that a single ecosystem can be created, where control is able to be exerted across all elements of an edge deployment. This means that the service can be optimised for different uses, enhancing quality of service as a result.

Design Considerations for Device Management at the Edge

Device management at the edge allows efficiencies of IoT deployments to be fully realised. By reducing the load on the cloud and its associated device management and analytics engines by moving processing to the edge, the system as a whole becomes more efficient.

Support for device types

The device ecosystem is very diverse and therefore management of devices at the edge must be tailored to support different types, classes and configurations of IoT device. Some of these devices will have more computing power than others and the device management platform will have to understand which devices can operate independently and which will rely more heavily on edge processing or require centralised management in the cloud.

Context awareness

As IoT devices perform a wide range of functions at all times of day, the device management services at the edge will have to be aware of the device context- this means not applying updates at a time when an IoT device might be especially active, or in certain locations.

Multi-access edge computing

Some edge devices will use multiple networks to communicate with the cloud (as per the ETSI Multi-access Edge Computing model[2]). This means that device management needs to be aware of which network a device is using and what data it being passed across it. This may affect configuration management if some updates can only be performed over certain networks.

[2] www.etsi.org/mec

6.1.2 Security

The IoT, by its very nature is a distributed, complex network of devices. Edge computing pushes much of the logic and data storage for effective operation of the IoT closer to the end devices, and having security services also distributed at the edge offers the opportunity to improve security capabilities, as well as offering native security for new low latency applications.

The edge has an important role to play in data security. Many advanced tools and techniques can be applied to ensure that the edge contributes to the security of the overall IoT deployment. With a vast array of different equipment and devices connected to the IoT, security services at the edge can be used to comprehensively secure or even isolate complex industrial environments such as smart factories and buildings, as well as used to ensure that data privacy is maintained in applications handling personal data such as CCTV or automated licence plate readers.

Security at the IoT edge should be treated the same as any other secure environment, but there are new tools that can be used to ensure security at new edge and device levels. For example, using strong identity management for devices at the edge means that authentication is more straightforward, as does robust definition of edge processes to ensure that they remain secure.

A number of security issues can be addressed at the edge in an IoT environment:

- **Firmware and other updates**

Secure update of firmware and other device updates from the edge using public key certification or secure transmissions such as SSL ensure that firmware upgrades are carried out securely.

- **Data Authentication**

Authentication of data and updates at the edge is important to retain secure environments. Authentication is likely to be via a certificate-based system. Implementation of this will need careful consideration to prevent poor performance of edge processing and latency.

- **Access Control**

Identity and permissions management at the edge is important to ensure that access to data at the edge is managed securely. Granting data access to third parties means that full access control policies must be in place.

- **Prevention of Denial of Service attacks**

Analysis of the data flow from IoT devices to spot and prevent characteristics of DDoS attacks.

Benefits of the Edge to this Use Case

- The distributed nature of edge computing for IoT means that malicious attacks aimed at the network are harder to instigate as attacking single nodes will only have limited impacts. The edge also offers more processing power to prevent attacks such as DDoS in addition to the central core.

- The IoT edge offers a new way of securing IoT end-points which may not be running the most up to date firmware or operating system. Security services at the edge can be used to ensure that devices with a high risk profile can be more easily isolated or have their data actively intercepted and secured.
- Data and device provenance – as processing and data storage moves closer to the edge, then the origination point of data is better understood and can be recorded with greater confidence.
- Processing of authentication, identity and access management – although sometimes constrained, additional processing power at the edge can be used for robust security processes as well as customer applications, meaning that security processes can be applied to ever increasing volumes of data.

Design Considerations

Using the edge environment to enhance IoT security could improve trust for the service being offered, but there are a number of considerations that should be considered in any deployment.

- **Constrained resources**

Processing resources at the edge are constrained compared with centralised services, and so the same models for securing data and access at the edge may not be the most effective to apply at the edge – at the very least they will need to be re-configured to best make use of the resources available.

- **Appropriate policies**

IoT at the edge must have relevant policies available that are applicable to both IoT and edge use cases. These policies must be applied whenever data access is granted or updates are to be sent to devices. Failure to implement robust policies will mean that attacks are more likely to succeed.

- **Minimise attack vectors**

By focusing IoT activity at the edge through, for example, only selective data generation and collection, some attack vectors could be minimised. Holding and transporting less data makes both the system more secure but also frees up system resources for better security management and authentication. The location of data processing needs to be appropriate – there may be classes of IoT data which can only be analysed in the cloud or highly secured data centres for example.

6.2 IoT SERVICE ENABLEMENT

6.2.1 Priority Messaging

Much of the data generated by the IoT will be of low value – unexceptional status updates and low priority data. However, some data will be of great importance and needs to be prioritised to ensure it is acted upon rapidly. This 'critical data' is likely to be a very small percentage of the total volume generated, yet is the most important. The scope of priority messaging goes beyond just single applications, as these message types could be used to initiate a cascade of actions across different applications and devices.

Examples of priority messaging include:

Transportation - accident alert that needs to be sent to following vehicles to enable them to avoid collision.

Health & Safety - fire alarm linked to building evacuation.

Environmental - rainfall or pollution above maximum safe levels linked to remedial activities.

Security – unauthorised activity leading to automated security actions e.g. doors closing; terrorism response in immediate vicinity; drones flying into no-fly areas.

Industrial – failure of critical component required immediate shutdown of other systems; construction worker in unsafe location.

The edge enables high priority data needs to be generated, sent, processed, and actioned more quickly than sending the data to the cloud.

Design Considerations

- **Fast processing at the edge** – low latency means that priority messages can be acted upon more quickly at the edge. Having relevant data storage and applications in a local cloudlet means that messages are received and acted upon quickly, without having to rely on centrally held data or applications.
- **Message association** – devices do not need to operate in isolation at the edge, and a priority message from one device may well be replicated by another nearby, meaning that the scale of any issue can be quickly judged.
- **Routing** – processing of priority messages at the edge means that their routing can be optimised through the rest of the network architecture, so they get to a specific endpoint in the fastest possible time.
- **Battery life** - data prioritisation at the edge means that low powered IoT devices can save battery life and processing power by leaving the actioning of critical data to the edge node.

Design Considerations

Priority messaging will need be considered as part of the overall design of IoT products, networks and data processing services. There are different types of priority messaging. Primarily, priority messages are in a known format as an output of a known process, for example, a fire alarm. Sometimes the message may not be identified as high priority by the device, but processing at the edge could still identify it as such. For example, voltage fluctuations in a smart grid may need the aggregation of data from different devices to determine priority.

High priority messaging and subsequent actions are enabled by the IoT edge. By only communicating with local IoT gateways and cloudlets, and keeping the impact of priority messaging local, faster responses can be assured. Data that has a wider impact beyond the immediate locale can be uploaded to the central cloud for further dissemination and decision making either immediately or after the local situation has been handled, for example, to feed in to larger datasets for analytics purposes.

For high priority messaging to be actioned quickly, near real-time processing of data is required. When tagged data is received, it can be easily identified and prioritised if the correct classifications are set-up within applications at the edge gateway. For other data, further processing could be used to identify data coming from multiple applications and devices to allow it to be prioritised in the same way.

The application on the edge node or gateway will need to make an automated decision as to what action to take with a high priority message. It may well be that a pre-defined process is initiated for known data types from known applications or devices. For example, a drone drifting into unauthorised areas can be redirected to a new flight path through a pre-defined process that all drones will need to recognise.

> **“By only communicating with local IoT gateways and cloudlets, and keeping the impact of priority messaging local, faster responses can be assured”**

6.2.2 Data Aggregation

As more IoT devices are connected, and more data generated, so it is likely there will be more replication of data from those devices. This could be multiple temperature readings from sensors in the same vicinity, or multiple vehicles reporting that they are stuck in the same traffic jam. Not all of this data needs to be sent back to centralised services, and the edge therefore has a role in either selecting which data to send or aggregating common data from multiple sensors together.

Examples of data aggregation:

- Data from multiple temperature sensors in the same location can be aggregated to produce statistical measures (min, max, mean etc).
- Traffic data derived from multiple vehicles in the same queue.
- Power outage reports from meters sending last gasp communications.
- Positive status reports from widespread connected equipment such as streetlights.

Benefits of Edge for this Use Case

There are several benefits to aggregating IoT data at the edge, before sending it onto the core.

- **Network efficiencies**

Data aggregation can create significant efficiencies in the IoT network. Aggregation can remove the need for replicating data across multiple systems, and performing the same processing multiple times on different systems. This means that there is no need to backhaul masses of replicated data, and therefore resources for data analytics can be used more effectively and data storage needs can be lowered. All of this means that the load on the core infrastructure is significantly reduced.

- **Latency improvements**

By having less data to sift through, quicker decisions could be made, so appropriate actions can be taken faster. By reducing the amount of data to be communicated and processed, latency should be improved.

- **Richer data sets**

Aggregated data provides valuable data sets where much of the data pre-processing has already been completed. This could aid machine learning in making more reliable predictions and allows patterns and trends to be more readily identified.

Design Considerations for Data Aggregation at the Edge

A large volume of similar data can overwhelm a network if sent at once, for example if a power outage results in every smart meter sending an alert at the same time. The role of the application running at the edge in these scenarios is to identify and aggregate multiple similar messages into just a handful of messages. There are a couple of ways of doing this:

- Aggregated data generation

If multiple messages are received by an edge node over a short period, these messages can be aggregated and the original messages deleted. The edge would not send on all messages or data when multiple messages are received. Instead, a new dataset is generated which summarises the data received. This would contain an overview of all messages received, potentially with periodic updates, until there is a status change.

- Sampling

A second option is to undertake sampling of devices. If a large volume of similar messages are received, the edge could elect to only monitor a handful of affected devices, rather than the whole fleet until the status of these changes. If cars enter a traffic jam, rather than taking status updates from every vehicle in the jam, data could only be taken from every 10th vehicle until they start to move again.

There are some challenges with data aggregation that need to be considered. For example, aggregation at the edge node has no benefit to the utilisation of the radio access channels since devices still send duplicate data over the radio network.

Also, aggregation may mean that the latest or most comprehensive data is not used for analysis, and so this should be considered according to the use case. In some situations, data timeliness may be more important than data aggregation, with out of date data being of little value.

Additionally, aggregating data also means that some data could be overlooked, which is particularly important if an application needs large volumes of data or a more complete current or historical data set for analysis.

6.2.3 Data Replication

Many IoT services are not localised – they are spread across a large geographic area or devices need to move between different locations on the mobile network. To support low latency at the edge some applications may need access to a localised data store at each location that they move to. This data store needs to be consistent across all instances of the application to ensure consistent results. This data may be time sensitive – a weather forecast, traffic conditions or distributed ledger instance for example. Without access to this data locally, applications will need to query the cloud, which will affect latency and creates single points of dependency. By replicating data across multiple sites, this issue can be avoided, and a seamless experience created at every location.

Benefits of Edge for this Use Case

Replicated data has the following benefits:

- Efficiencies of IoT at the edge can be fully realised, across multiple locations. Not having to call to the cloud for master data significantly increases the performance of an application.
- Low latency support – by having local direct access to relevant data at the edge, low latency can be better realised.
- Disaster recovery - by distributing data across multiple locations, disaster recovery becomes possible, as data can be replicated across multiple locations, reducing the risk of data loss even if a node were to fail.
- Scalability - heavy processing of data can be distributed across all available processing power, meaning that a system is not reliant on a single cloud or core source for transactional processing power. Distribution may be possible across different vendor systems with the correct standards in place.

Design Considerations for Replicated Data

There will generally be a need to have certain data ultimately held in the cloud for long-term storage, analytics and future processing. However, real-time data and logic that is needed for the operation of a wide variety of services under an IoT edge computing architecture must be available as an application needs it, without the latency implicit in storing and processing data centrally.

- **Time synchronisation**

In order to synchronise data replicated across multiple locations, there will likely need to be time synchronisation at all edge locations. To ensure consistent data across multiple nodes, time must be accurate, otherwise data at different locations may not be fully synchronised.

- **Preventing task duplication**

The use of distributed ledgers to store transactions at edge nodes is a good way to ensure consistency of data and messages across replicated nodes – once a message or task has been completed, this could be recorded in the ledger to ensure that the activity is replicated to every peer node.

Master data

Although distributed data is important for real-time transactions, a master data set is likely to be needed from a central location, to govern how the local instances should operate. Centralised operation of multiple cloudlets is needed to ensure that there is a distinct hierarchy in place for issue resolution and effective control of multiple systems.

Personal data

As data is replicated across nodes, special consideration should be given to personal data that would be affected by the EU GDPR and similar regulations. Data replication processes should have scope to identify personal data and the way that should be handled.

Node to node communication

If enabled, nodes are able to communicate with each other. This means that a process of ensuring that data sets are replicated without reference to master data sources inside the cloud. Data must be replicated directly into other nodes where the local application may need access to up to date data.

Security

Data authentication at the edge requires local attributes to be held to ensure data is authenticated with low latency. Holding this information in the centralised cloud would slow down operations where authentication is needed.

Device management

Device configurations may need to be replicated across multiple nodes at the edge to ensure that any device management services are applied at the appropriate time or location.

6.2.4 Cloud Enablement

The edge is expected to be attractive to cloud vendors as it offers them the potential to better distribute data storage and processing power to support low latency services, support higher systems availability and also to reduce the load on their existing data centres. A mutually beneficial relationship between cloud providers and mobile operators should exist at the edge, where operators have local resources to enable IoT edge for cloud providers, and cloud providers have platforms suitable for enabling a wide range of services on the operator's edge infrastructure. Local infrastructure also meets customer and regulatory requirements about data storage on premises or in the country, if edge resources can be located in the relevant location. Some IoT users such as factories or smart cities may also insist on their data being stored and managed locally.

From a service point of view, IoT edge environments will rarely operate completely in isolation. A connection to a centralised cloud will often be required for control, monitoring and update purposes at the very least, and so the dynamic between the edge and cloud is a complex one. A hybrid edge and cloud architecture can offer the best of both worlds.

Benefits of Edge for this Use Case

Linking the IoT edge and cloud has many benefits, across network, application and data management.

- **Maximise use of resources both locally and centrally**

By linking the edge and the cloud, the most appropriate resources for specific tasks can be identified and allocated. This means efficient usage of resources in the operators domain, perhaps held at local base stations, or on the device itself, or in the cloud providers domain, where huge data volumes can be centrally stored.

- **Integrated view of data**

Integrating the edge and cloud means that users will have a view over the status and location of all data relevant to their application or service. A seamless view means that quality of service can be achieved without having to resort to only working at the cloud or at the edge.

- **Built-in scalability**

By enabling distributed resources, scalability of IoT services at the larger end of the scale becomes achievable. Even if resources at the edge are unable to cope with the scale of operations required by a deployment, there can be a fallback to the cloud, meaning that some or all of the service can be maintained.

- **Security**

Full integration of edge and cloud means that data security can be overseen from a single source. Utilising the relevant cloud agents ensures that data can be securely transmitted across the cloud, edge and device.

- **Enablement of new business models**

Business models including Infrastructure as a Service (IaaS) or higher response level Service Level Agreements enabled by lower latency can be introduced with integrated cloud and edge access. Unified billing for processing and storage can be managed across the cloud and edge in processes seamless to the end user. The edge can reside in different locations of the IoT deployment chain, and any cost benefits can be used to create new business models.

Design Considerations

Approach to data processing

A "Slow Lane" / "Fast Lane" approach is necessary to ensure that data is held in the correct location and processing resources are prioritised accordingly. "Fast Lane" data will obviously be dealt with first, and likely to be processed as fully at the edge as possible, for all the reasons stated elsewhere in this paper. "Slow Lane" data on the other hand is likely to be backhauled at an appropriate time to the cloud for processing. "Slow Lane" data may provide a more comprehensive view of what is occurring, but it will take time to achieve that view. To facilitate this model appropriate data analytics models and data categorisation need to be put into place by the analytics provider.

Support for cloud agents

Many cloud providers have dedicated edge agents that will manage the relationship between the edge and the cloud. Recognition of these and integration of support for them into wider application and device management platforms will enable cloud integration for a wide range of service providers and encourage portability of applications.

Developer support

Developers commonly look to incorporate cloud support and functionality in their applications, and the same will be true of IoT applications at the edge. Access to the edge infrastructure and platforms should be implemented in the relevant tools to enable the developer community to take advantage of benefits edge offers the IoT.

> **"Integrating the edge and cloud means that users will have a view over the status and location of all data relevant to their application or service."**

6.3 IoT SOLUTIONS

6.3.1 IoT Image and Audio Processing

Devices such as cameras, including CCTV, and microphones can provide data for processing by IoT platforms and applications, such as licence plate reading or monitoring noise pollution. IoT edge introduces new ways of analysing this data without having to backhaul the entire image or audio stream. An edge cloudlet can be used to process the image, video or audio data to determine key information, such as licence plate numbers or the number of people in an area, meaning that only a small amount of data, such as the licence number itself is forwarded or stored. Other examples where the camera can be used as a sensor include for monitoring of environmental conditions such as river levels, monitoring crowd density, or in industrial IoT, whether on a factory floor, monitoring powerlines from a drone, or listening for flow through pipelines to identify leaks.

Benefits of Edge for this Use Case

- **Low cost**

Cameras and microphones are relatively low cost to procure, install and maintain for the insights they can provide; use of edge processing means that network management costs are also managed effectively, making them an attractive general purpose alternative to dedicated IoT sensors.

- **Significant reduction in network backhaul**

By identifying objects within images, without needing to send the image itself to upstream servers, the amount of data that needs to be transmitted back to the core is significantly reduced.

- **Quick decision making**

Fast processing means that it is possible to support a wider range of real-time or near real-time applications – speeding up the management of production lines or enabling new ways of charging drivers at tollgates and so on.

- **More flexible IoT sensor arrays**

By adding camera data to IoT deployments, a more comprehensive analysis can be taken, as cameras can add more general context (through both imaging of a location and broad image coverage) than many other types of IoT sensors.

- **Enabling new use cases**

New IoT use cases become possible with the use of cameras and microphones as sensors. For example, the use of image processing for recognising yields, pests and diseases of crops whilst they grow.

Design Considerations

Design and deployment considerations for image and audio processing mainly focus on image and camera setup to ensure that the data source is good enough for analysis, but other considerations are also relevant.

Image or audio quality

Information can only be recognised from a camera image or audio stream if the quality is good enough – this means high enough resolution, but also the ability to recognise from an image or audio stream in all environmental conditions – day and night, sun and rain, crowd/ traffic. If this is not possible the analytics will not be fit for purpose.

Image & audio format

Images can be still or video and along with audio can come in a variety of formats and obtained using a variety of standard and proprietary protocols. Analytics engines will need the image data to be decoded to frame data. Video images may need to be broken down into a series of still images for analysis. Collected data will ideally be in a common, open format so that is can be managed by a range of analytics engines.

Data analytics

A picture is worth a thousand words, but for IoT analytics purposes, there needs to be a very clear definition of the parameters that are needed from an image or audio file, so that machine learning processes can be trained. This means that a clear definition of the image topography or audio landscape and how it relates to the data to be extracted must be defined in advance.

Camera setup

In some cases, it may be simpler if the camera is in a fixed position with a number of reference points that do not change, so that the image processing engine can accurately identify the area of the image to process.

7. Operator Opportunity and Potential Next Steps

7.1 OPERATOR OPPORTUNITY

IoT Foundation

The edge is a natural evolution of today's IoT architecture deployed by mobile operators. Evolution of platforms, processes and propositions will enable operators to introduce edge seamlessly into their foundational IoT services, such as connectivity and device management, whilst retaining the attributes which they are renowned for – connectivity, security and scalability.

The edge offers new ways of creating efficient connectivity services specifically designed to benefit IoT deployments by reducing the amount of data which is backhauled to the cloud. Better options for managing devices in the field become available, with distributed management of firmware updates and applications possible. IoT platforms will need to extend their reach to the edge, and operators will be in a strong position to ensure that network, security and cloud services are integrated effectively at the edge.

IoT Service Enablers

Data Management at the edge will be crucial for operators in ensuring quality of service and effective analytics of IoT data. Managing customer data in a secure fashion and ensuring a seamless integration between user, cloud, application, edge and device will allow a new generation of service offerings from operators.

IoT devices can benefit greatly from local processing power, data management and analytics. Having the ability to prioritise messages and manage the large volume of data from massive IoT deployments means that operators can both maximise the cost savings on their own network, but pass on better service levels to their customers.

There is a mutually beneficial relationship with cloud providers that operators can also build, through a combination of local infrastructure controlled by the operator, and the data management platforms from cloud providers to enable a new generation of IoT management services.

IoT Solutions

The edge offers operators both opportunities for new IoT solutions, and new ways of managing existing services. IoT sensors which generate large amounts of data such as cameras can be used to provide new or better insights, and data from multiple sensors can be combined to create new levels of analysis.

Advanced IoT solutions such as support for V2X communications or smart factories can benefit from these new opportunities that operators can introduce. By making the edge the standard place for managing applications and services, new levels of service management, automation and precision can be achieved.

7.2 POTENTIAL NEXT STEPS

Deployment of IoT edge and applications that will utilise it is not a trivial task, and there are a number of challenges that will need to be addressed. Operators will need to strike a balance between the benefits that edge brings in scaling and managing IoT deployments and the costs of setting up a service. A number of potential next steps for the industry have been identified:

Common Framework

- Define end user requirements for customers and application developers.
- Define where the IoT edge resides for defined use cases – at the cloudlet in the data centre, on the base station, or closer to the device.
- Understand deployment and business models and match to the infrastructure that could be utilised.

Evaluate Solutions

- Evaluate different edge models, consortiums and technologies for their suitability to IoT deployments.
- Understand the need for IoT platform extensions at the edge, and how the functionality available maps to IoT use cases.

Operator Roles

- Understand the roles of different partners through the value chain and how the operator can create value for them.
- Understand cloud offerings at the edge and how the operator can integrate with them or support them. Engage with cloud providers to create a mutually beneficial model for deployment of edge services from both operator and cloud provide.
- Undertake relevant pilots and other activities to investigate the benefits of edge for IoT customers.

8. Conclusions

IoT edge offers significant potential benefits for mobile operators, their customers and their partners. There are distinct differences between the use cases for IoT edge and other edge computing use cases. IoT tends to be focussed very much on managing a high volume of devices and the data that they generate, to ensure fast data processing and reduce the operational complexity of managing such a large fleet of mission-critical devices.

Operators are naturally well placed to offer edge services and have existing strengths which should prove to be a strong attribute in future discussions with application developers and cloud providers who will both need access to infrastructure at the edge to enable their own business models. Adding capabilities as outlined in this document will mean that operators have more control over the edge and deployment of IoT services using it. In turn, they will be able to build stronger relationships across the IoT ecosystem. However, the cost and complexity of edge implementation will be a consideration.

Many vertical services can benefit from IoT edge – smart factories, intelligent transport, smart cities and energy companies can all benefit from data processing and storage at the edge. Once these industries are able to deploy their first services using dedicated IoT tools and applications at the edge, new opportunities for new services will open up, and integration between different verticals will be possible – electric cars linked to the smart grid or intelligent buses being routed to where there are large groups of people waiting will all be possible with the low latency and fast data processing capabilities at the edge.

The edge for IoT is relatively immature and the operator-led tasks of decentralising existing IoT deployments, and re-aligning platforms and infrastructure to enable new IoT services at the edge is not trivial. However, operators are well placed to manage the required infrastructure and integrate it into their own existing networks. Mobile networks are already well suited to IoT edge, as they are highly distributed and match the footprint of many IoT deployments, and this is particularly the case where operators are rolling out Mobile IoT solutions. Operators understand the needs of IoT customers and how they can combine both high quality connectivity with fast data processing, leading to AI driven automation, enabling the next generation of IoT services on mobile networks.

For more information please visit:
www.gsma.com/IoT

GSMA HEAD OFFICE
Floor 2
The Walbrook Building
25 Walbrook
London EC4N 8AF
United Kingdom
Tel: +44 (0)20 7356 0600
Fax: +44 (0)20 7356 0601