

## **Draft new Recommendation ITU-T Y.IMT2020-CEFEC**

### **Framework of edge computing capability exposure for IMT-2020 networks and beyond**

#### **Summary**

There are various edge computing capabilities that can be exposed to applications. With the exposure of such capabilities, the applications, including edge computing applications and non-edge computing applications, are able to obtain augmented information from which these applications can benefit, especially, but not limited to, in terms of performance.

This Recommendation specifies the framework of edge computing capability exposure for IMT-2020 networks and beyond.

#### **Keywords**

Capability exposure; edge computing; framework; functional component; IMT-2020; reference point

## Table of Contents

1. Scope.....	3
2. References.....	3
3. Definitions.....	3
3.1. Terms defined elsewhere .....	3
3.2. Terms defined in this Recommendation .....	4
4. Abbreviations and acronyms.....	4
5. Conventions .....	4
6. Framework of edge computing capability exposure.....	4
7. Reference points of edge computing capability exposure .....	6
7.1. Reference point R1 .....	6
7.2. Reference point R2 .....	7
7.3. Reference point R3 .....	7
8. Procedures for edge computing capability exposure .....	7
8.1. Edge application capability exposure .....	7
8.2. Procedures of capability edge network capabilities exposure.....	10
8.3. Procedures of edge computing capability exposure considering capability classification .....	11
8.4. Procedures of edge computing capability exposure considering capability exposure cache.....	12
8.5. Procedures of edge computing capability exposure considering continuity needs .....	13
9. Considerations about charging of edge computing capability exposure .....	14
10. Security considerations .....	14
Bibliography.....	15

## Draft new Recommendation ITU-T Y.IMT2020-CEFEC

### Framework of edge computing capability exposure for IMT-2020 networks and beyond

#### 1. Scope

This Recommendation specifies the framework of edge computing capability exposure for IMT-2020 networks and beyond.

The following aspects are addressed:

- functional components of the framework and their interactions;
- reference points of the framework;
- procedures of edge computing capability exposure.

#### 2. References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- [ITU-T Y.3101] Recommendation ITU-T Y.3101 (2018), *Requirements of the IMT-2020 network*
- [ITU-T Y.3104] Recommendation ITU-T Y.3104 (2018), *Architecture of the IMT-2020 network*
- [ITU-T Y.3105] Recommendation ITU-T Y.3105 (2018), *Requirements of Capability Exposure in the IMT-2020 network*
- [ITU-T Y.3108] Recommendation ITU-T Y.3108 (2019), *Capability exposure function in the IMT-2020 networks*

#### 3. Definitions

##### 3.1. Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

- 3.1.1. IMT-2020** [b-ITU-T Y.3100]: Systems, system components, and related technologies that provide far more enhanced capabilities than those described in [b-ITU-R M.1645].

NOTE – [b-ITU-R M.1645] defines the framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000 for the radio access network.

- 3.1.2. third party (3rd party)** [b-ITU-T Y.3100]: In the context of IMT-2020, with respect to a given network operator and network end-users, an entity which consumes network capabilities and/or provides applications and/or services.

NOTE 1 – An example of 3rd party, a virtual network operator (VNO) may use capabilities exposed by a network operator, e.g., to manage specific network slices. Another example of 3rd party, a service and/or application provider (e.g., an OTT player) may provide applications and/or services to enhance the network capabilities.

NOTE 2 – Network end-users are not regarded as 3rd parties.

### **3.2. Terms defined in this Recommendation**

None.

### **4. Abbreviations and acronyms**

This Recommendation uses the following abbreviations and acronyms:

API	Application Programming Interface
ECA	Edge Computing Application
ECP	Edge Computing Platform
ECCEMP	Edge Computing Capability Exposure Management Platform
IMT-2020	International Mobile Telecommunication 2020
IP	Internet Protocol
IT	Information Technology
NCEP	Network Capability Exposure Platform
UE	User Equipment

### **5. Conventions**

None.

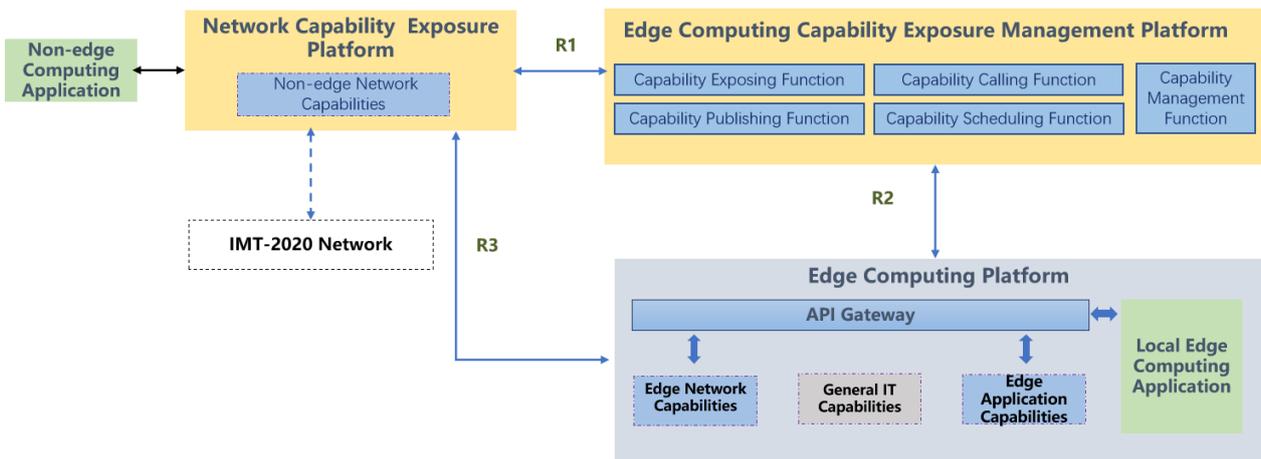
### **6. Framework of edge computing capability exposure**

According to [ITU-T Y.3105], the exposure of the IMT-2020 network capabilities will bring new opportunities to network operators, vendors and third parties. According to the architecture of IMT-2020 network [ITU-T Y.3104] and the functionalities of capability exposure function in the IMT-2020 networks [ITU-T Y.3108], network capabilities can be exposed to third party applications.

There are various edge computing capabilities that can be exposed to applications. With the exposure of such capabilities, the applications, including edge computing applications (ECA) and non-edge computing applications, are able to obtain augmented information from which these applications can benefit, especially, but not limited to, in terms of performance.

NOTE 1 - Non-edge computing applications are applications that are centrally deployed comparing with edge computing applications that are deployed at the edge.

Figure 6-1 illustrates the framework of edge computing capability exposure for IMT-2020 networks and beyond, including the interactions between the functional components of the framework, i.e., the Edge Computing Platform (ECP), the Edge Computing Capability Exposure Management Platform (ECCEMP) and the Network Capability Exposure Platform (NCEP).



**Figure 6-1 Framework of edge computing capability exposure for IMT-2020 networks and beyond**

The functional components of the framework are described in the following sub-clauses and the reference points, R1, R2 and R3, as shown in Figure 6-1, are described in Clause 7.

NOTE 2 - The non-edge network capabilities exposed by IMT2020 network, and the related exposure, are out of scope of this Recommendation – they are shown using dashed lines in Figure 6-1.

NOTE 3 - Figure 6-1 shows ECAs deployed on the local ECP (local edge computing applications), while ECAs may also be deployed on other (remote) ECPs (these ones are not shown in Figure 6-1 for simplicity).

## 6.1. Edge Computing Platform

This functional component provides three kinds of capabilities, i.e., edge network capabilities, edge application capabilities and general IT capabilities. The edge network capabilities and the edge application capabilities constitute the edge computing capabilities. These edge computing capabilities provide services in the form of Application Programming Interface (APIs). The ECP is used to provide the required environment for ECAs and edge computing capabilities. Different ECPs may provide different sets of edge computing capabilities and one ECA (residing on a given ECP) may request the support of capabilities residing in different ECPs.

In particular, concerning the three kinds of capabilities:

- the edge network capabilities are provided by network operators, and provide network status information, such as bandwidth, latency, etc.
- the edge application capabilities are provided by ECAs, and include, but are not limited to, video decoding/encoding and image processing.
- the general information technology (IT) capabilities are capabilities internal to the platform itself, including but not limited to, API management capabilities. These capabilities are not exposed.

The API gateway on the ECP is generally used as a unified call entry by the ECAs.

NOTE – Multiple ECPs can interact with the ECCEMP and the NCEP.

## 6.2. Edge Computing Capability Exposure Management Platform

This functional component is used to provide unified edge computing capability management and operations, including, but not limited to, capability exposure, capability publishing, capability calling, capability scheduling and capability management.

There are five key functions in the ECCEMP:

- **Capability Exposure:** the ECCEMP exposes edge computing capabilities deployed in the ECP, including the capabilities' identification and location.
- **Capability Publishing:** the ECCEMP publishes the capabilities to the NCEP, with capability related information, such as capability s' identification and location.
- **Capability Calling:** the ECCEMP receives the ECA's capability discovery request, or the non-edge computing application's capability discovery request from the NCEP, and, after the authentication of the request, returns the assigned capability related information to ECA or non-edge computing application.
- **Capability Scheduling:** the ECCEMP configures and enforces the capability scheduling policy with respect to a given capability. NOTE 1 – For example, it can configure the traffic congestion control policy, with candidate capability 1 and candidate capability 2 with respect to the same capability, and, if candidate capability 1 is busy, the platform chooses candidate capability 2 to provide services.
- **Capability Management:** the ECCEMP's management capabilities include status monitoring, usage analysis and handling of capabilities. NOTE 2 – An example of such management capabilities is the off-lining of a capability which returns error frequently and the related notification to the corresponding ECA.

NOTE 3 – From an implementation perspective, the ECCEMP may be integrated in a broader management system, e.g., an edge computing management system.

### **6.3. Network Capability Exposure Platform**

This functional component acts as the capability calling portal of non-edge computing applications, supporting the edge computing capability calling to the ECP through the ECCEMP. It receives the capability publishing information from the ECCEMP, and checks and stores the capability information.

The NCEP also hosts the non-edge network capabilities exposed by IMT2020 network, and receives non-edge network capability call requests from the ECP.

## **7. Reference points of edge computing capability exposure**

The reference points of capability exposure function of edge computing capabilities are addressed in this clause.

### **7.1. Reference point R1**

R1 is the reference point between the NCEP and the ECCEMP. This reference point is used to deliver:

- the edge network capabilities and the edge application capabilities exposure related information to the NCEP.
- the edge network capabilities and the edge application capabilities management related information to the NCEP, including but not limited to, capability publishing information, capability deletion information and capability modification information.

NOTE 1 – R1 is also used to deliver the non-edge network capability call requests of ECA. This is out of scope of this Recommendation.

NOTE 2 - An example of edge network capabilities exposure information is the edge bandwidth management capability exposure information. Via the exposure of this edge network capability, the NCEP may request, via the ECCEMP, bandwidth related operations including creation, cancellation, modification and query, concerning a specific ECA running over a given ECP.

NOTE 3 - An example of edge application capabilities exposure information is the edge application user management capability exposure information. Via the exposure of this edge application capability, the

NCEP may request, via the ECCEMP, user management operations including user access control, concerning a specific ECA running over a given ECP.

## **7.2. Reference point R2**

R2 is the reference point between the ECCEMP and the ECP. This reference point is used to deliver:

- the edge computing capability management related information, such as information on ECA bandwidth management, ECA user management and dynamic capability release management,
- the real-time status information of edge computing capabilities, such as load, response latency etc.,
- the edge computing capability calling requests and the edge computing capability scheduling results,
- the information on the use of edge computing capabilities, such as start time and end time.

## **7.3. Reference point R3**

R3 is the reference point between the ECP and the NCEP. This reference point is used to deliver the edge computing capability requests from non-edge computing applications to the ECP.

NOTE - R3 is also used to deliver the ECA's capability requests about network capabilities to the NCEP. This is out of scope of this Recommendation.

## **8. Procedures for edge computing capability exposure**

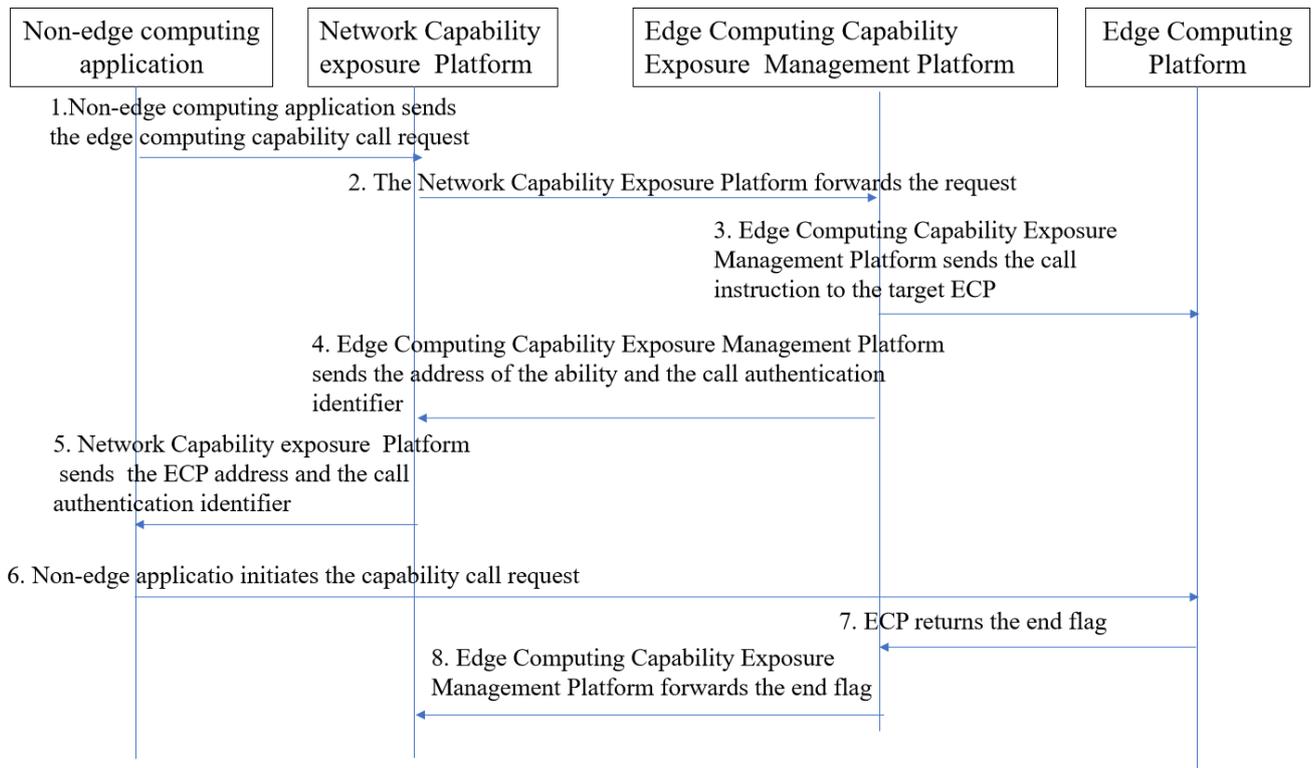
This clause describes in sub clauses 8.1 and 8.2 the procedures of the two kinds of edge computing capabilities exposure, i.e., the edge network capabilities exposure and the edge application capabilities exposure. It then describes in sub clauses 8.3, 8.4 and 8.5 the general procedures of edge computing capability exposure considering three different factors, i.e., capability classification, capability cache and continuity needs.

### **8.1. Edge application capability exposure**

#### **8.1.1. For non-edge computing applications**

The following describes the procedures of edge application capability exposure for non-edge computing applications.

Figure 8-1 illustrates the procedures of edge application capabilities exposure for non-edge computing applications.



**Figure 8-1 Procedures of edge application capabilities exposure for non-edge computing applications**

When a new capability is launched on the ECP, it is released to ECCEMP in real time through R2. Through R1, ECCEMP registers the new available capabilities to the NCEP. In addition, it maintains the location of ECP and the list of capabilities deployed on ECP. The ECCEMP synchronizes the capability call template with the NCEP.

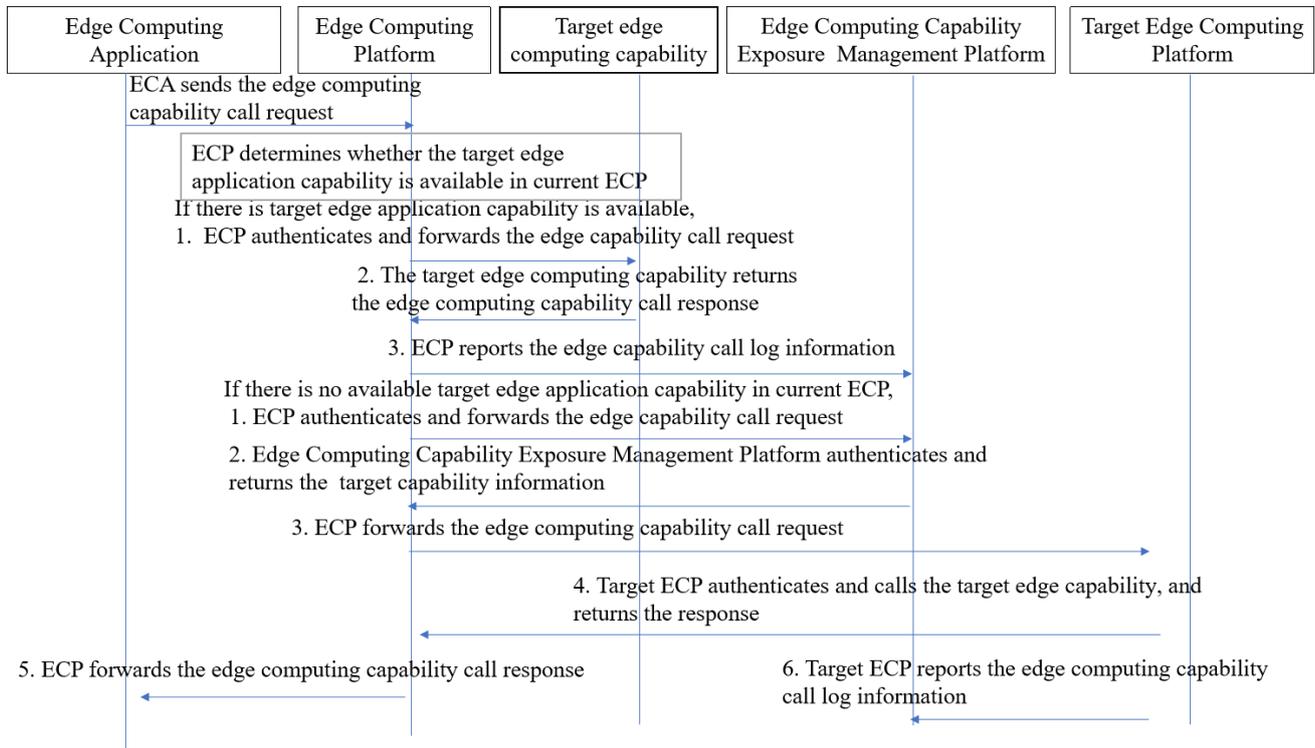
The procedures are as follows:

1. The non-edge computing application sends the edge computing capability call request to the NCEP, carrying the location information of the non-edge computing application.
2. The NCEP forwards the request carrying application location information to ECCEMP.
3. According to the location information of the non-edge computing application, the ECCEMP sends the call instruction of the selected capability including but not limited to call authentication identifier to the target ECP, considering the deployment location, availability, and load conditions of edge computing capabilities.
4. The ECCEMP returns the ECP's Internet protocol (IP) address and the call authentication identifier of the selected capability to the NCEP.
5. The NCEP continues to return the ECP's IP address that provides the capability and the call authentication identifier to the non-edge computing application.
6. The non-edge computing application carrying the call authentication identifier initiates the capability call request that conforms to the calling template to the target ECP.
7. After the transmission is completed, ECP returns the end flag (i.e., an edge computing capability call is completed) to ECCEMP.
8. The ECCEMP forwards the end flag to the NCEP.

### 8.1.2. For edge computing applications

The following describes the procedures of edge application capability exposure for ECA.

Figure 8-2 illustrates the procedures of edge application capabilities exposure for ECA.



**Figure 8-2 Procedures of edge application capabilities exposure for ECA**

ECA sends the edge computing capability call request to ECP with the authentication information, and ECP determines whether the target network capability is available in current ECP.

If the target edge computing capability is available in current ECP, the procedures are as follows:

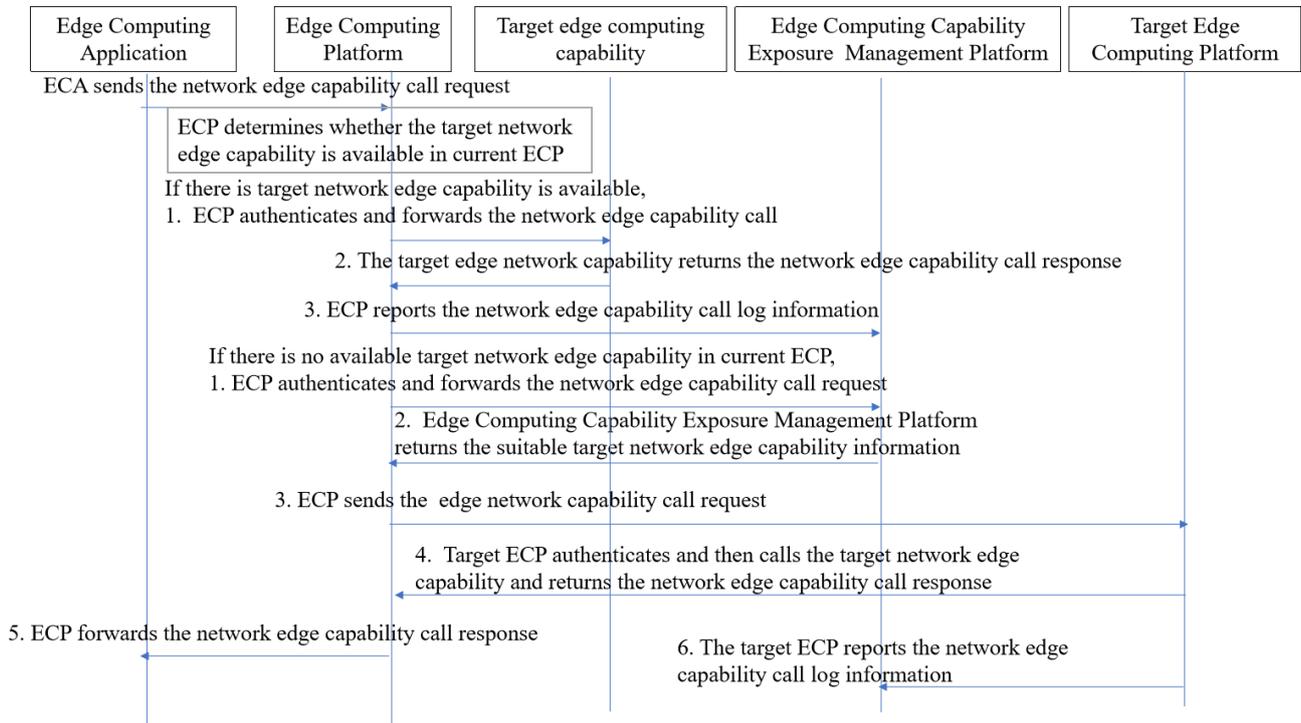
1. ECP authenticates the edge computing capability call request and then forwards the edge computing capability call request to the target edge computing capability.
2. The target edge computing capability returns the edge computing capability call response.
3. ECP reports the edge computing capability call log information to the ECCEMP.

If there is no available target edge computing capability in current ECP, the procedures are as follows:

1. ECP authenticates the edge computing capability call request and forwards the edge computing capability call request to the ECCEMP, with the authentication information.
2. The ECCEMP authenticates the edge computing capability call request, and returns the suitable target ECP's IP address information and authentication information.
3. ECP forwards the edge computing capability call request to target ECP with authentication information.
4. Target ECP authenticates the edge computing capability call request, and then calls the target edge computing capability and returns the edge computing capability call response.
5. ECP forwards the edge computing capability call response to ECA.
6. Target ECP reports the edge computing capability call log information to the ECCEMP.

## 8.2. Procedures of edge network capability exposure

Figure 8-3 illustrates the procedures of edge network capability exposure.



**Figure 8-3 Procedures of edge network capability exposure**

ECA sends the edge network capability call request to ECP with the authentication information, and ECP determines whether the target edge network capability is available in current ECP.

If the target edge network capability is available in current ECP, the procedures are as follows:

1. ECP authenticates the edge network capability call request and forwards the edge network capability call request to the target edge network capability.
2. The target edge network capability returns the edge network capability call response.
3. ECP reports the edge network capability call log information to the ECCEMP.

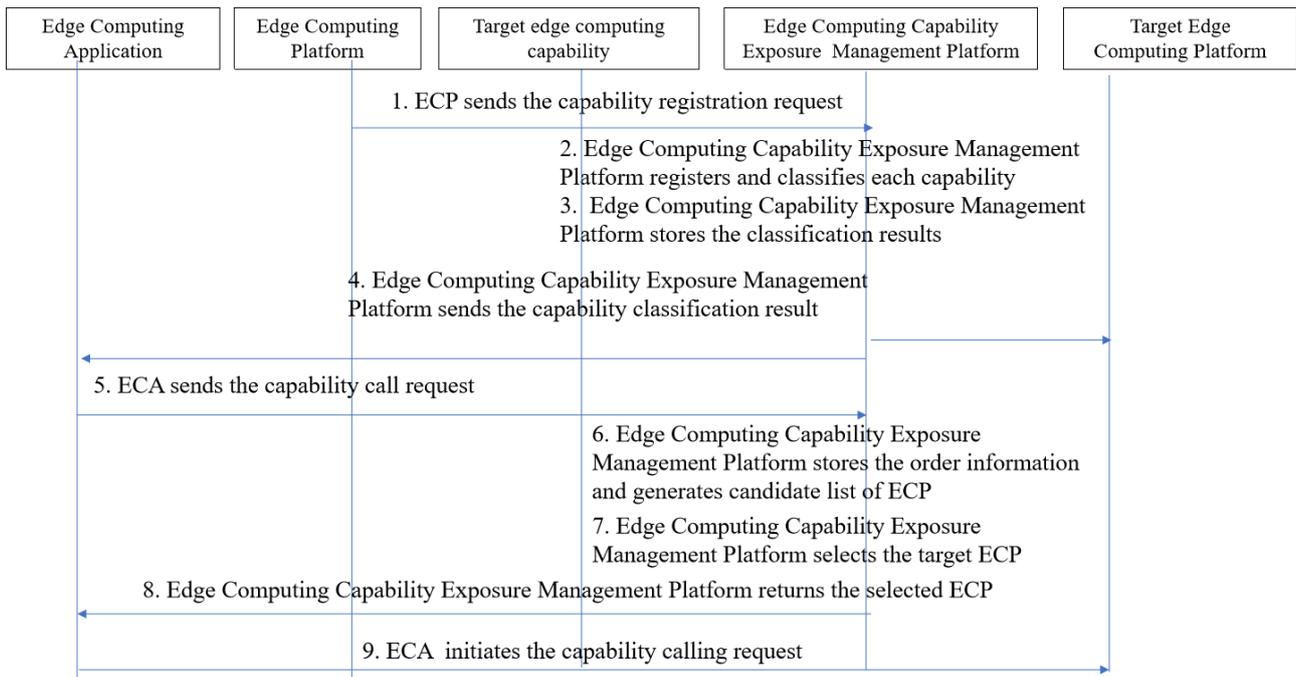
If there is no target edge network capability in current ECP, the procedures are as follows:

1. ECP authenticates the edge network capability call request and sends the edge network capability query request to the ECCEMP.
2. The ECCEMP returns the target edge network capability information, including the target ECP and authentication information.
3. ECP sends the edge network capability call request to the target ECP, with the authentication information.
4. The target ECP authenticates the edge network capability call request, and then calls the target edge network capability and returns the edge network capability call response.
5. ECP forwards the edge network capability call response to ECA.
6. After finishing the edge network capability calling, the target ECP reports the edge network capability calling log information to the ECCEMP.

### 8.3. Procedures of edge computing capability exposure considering capability classification

For the same edge computing capabilities deployed at multiple ECPs, the edge computing capabilities may have different performance levels according to multiple parameters. As an example, for edge network capabilities, the parameters include capability performance, capability access capacity and capability response delay. Also, the edge computing capability calling request may require different performance levels of edge computing capability, so the evaluation and classification of edge computing capabilities is important to assist the edge computing capability scheduling to improve the edge computing capability exposure performance.

Figure 8-4 illustrates the procedures of edge computing capability exposure considering capability classification.



**Figure 8-4 Procedures of edge computing capability exposure considering capability classification**

The procedures are as follows:

1. ECP sends the capability registration request to the ECCEMP.
2. The ECCEMP registers the capabilities and classifies them based on the capability related parameters including but not limited to capability performance, capability access capacity, capability response delay.
3. The ECCEMP stores the classification results.
4. The ECCEMP sends the capability classification results to the ECP which deployed the capability, and also to the ECAs which ordered and called the capability. NOTE 1 – ECAs may order specific capabilities to the ECCEMP. The ECCEMP sends the capability classification results only to the ECAs that ordered specific capabilities.
5. ECA sends the capability calling request to the ECCEMP, which carries the capability order information and the corresponding capability classification requirements. NOTE 2 – The capability order information includes, but is not limited to, the ordered capability identification and also the ordered capability classification information.
6. The ECCEMP stores the capability order information and generates the list of candidate ECPs according to the ECA’s capability calling request.

NOTE 3 - Each candidate ECP in the list has the capability corresponding to the capability classification requirement.

7. The ECCEMP selects the target ECP according to real-time performances, including but not limited to resource occupancy rate, response delay, upstream rate, downstream rate.

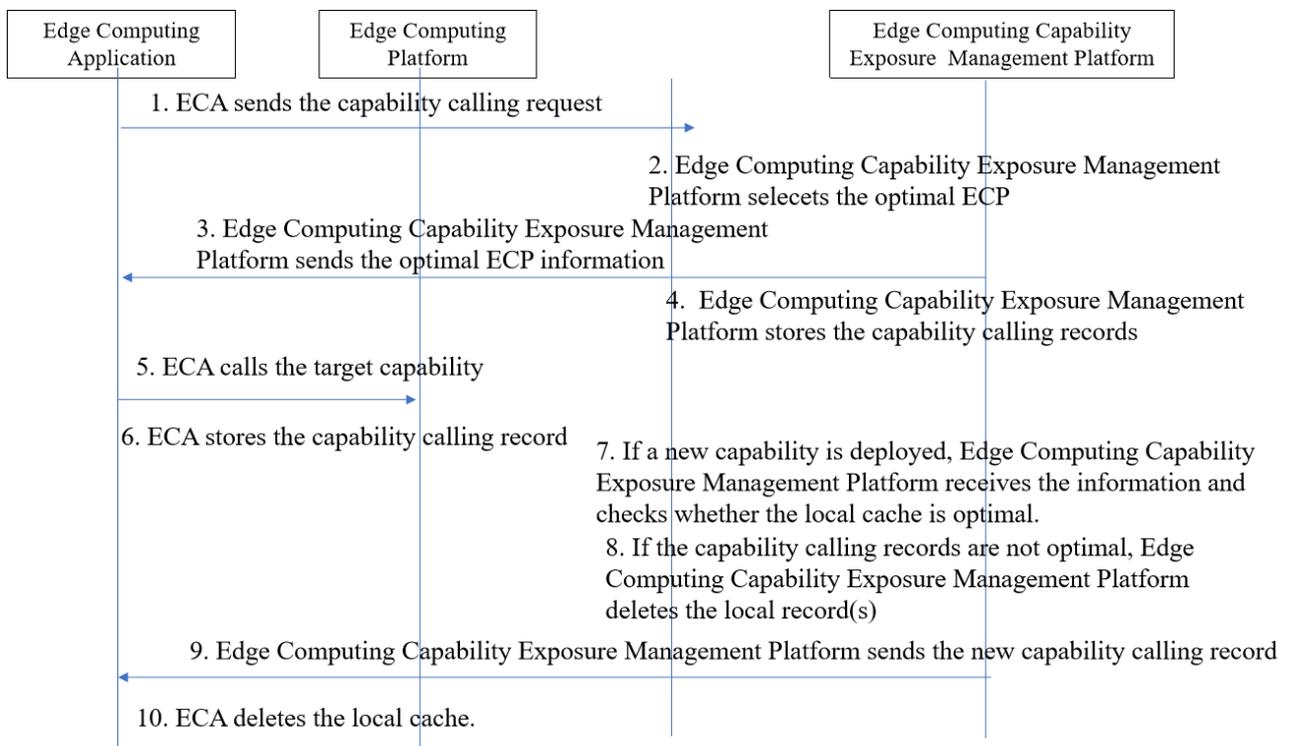
8. The ECCEMP sends the capability response to ECA, indicating the target ECP.

9. ECA initiates the capability calling request to the target ECP.

#### 8.4. Procedures of edge computing capability exposure considering capability exposure cache

The ECCEMP generates the capability exposure cache according to the capability calling record for further capability calling. ECA also stores the capability calling record and generates a local cache. The capability exposure cache is able to reduce the response time of capability calling process.

Figure 8-5 illustrates the procedures of edge computing capability exposure considering capability exposure cache.



**Figure 8-5 Procedures of edge computing capability exposure considering capability exposure cache**

The procedures are as follows:

1. ECA initiates the capability calling request to the ECCEMP.
2. The ECCEMP receives the request and selects the optimal ECP which deployed the target capability.
3. The ECCEMP sends the optimal ECP information to ECA.
4. The ECCEMP stores the ECA capability calling record.
5. ECA receives the optimal ECP information and calls the target capability.

6. ECA stores the capability calling record on ECP.

7. If a new capability is deployed, the ECCEMP receives the information and checks whether the local cache of capability calling records is optimal.

If the local cache of capability calling records is not optimal, the following procedures apply:

8. The ECCEMP deletes the local record(s) of the corresponding capability.

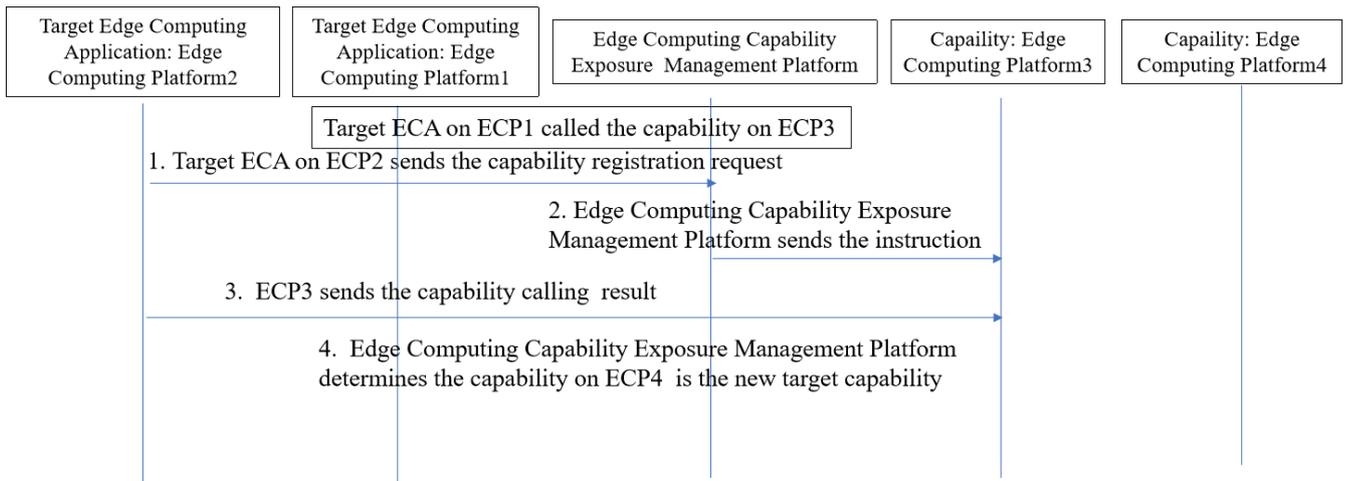
9. The ECCEMP sends the new capability calling record(s) to ECA.

10. ECA deletes the local cache. And when ECA requests again that capability, it sends the capability calling request to the ECCEMP.

### 8.5. Procedures of edge computing capability exposure considering continuity needs

For edge computing capability exposure, when a user equipment (UE) that initiated the capability call moves to a new location, and this movement implies that the previous capability is not optimal, the capability relocation needs to be triggered. Consequently, the capability continuity needs of UE have to be considered.

Figure 8-6 illustrates the procedures of edge computing capability exposure considering continuity needs.



**Figure 8-6 Procedures of edge computing capability exposure considering continuity needs**

As an example, before UE moves, the UE associated ECA (deployed on ECP1) called the target capability on ECP3. When UE moves to a new location, the UE associated ECA may be switched from ECP1 to ECP2. As the target capability on ECP3 may not be optimal anymore after the UE movement, the new target capability on ECP4 will be selected according to the following procedures.

According to the example described above, the procedures are as follows:

1. When the UE associated ECA is switched from ECP1 to ECP2, the ECCEMP receives the capability calling request of UE associated ECA on ECP2, with the capability calling request providing the identifier of the UE associated ECA.

NOTE 1 - The capability calling request includes the identification information of the UE associated ECA on ECP2 and, the capability continuity needs.

2. For the ECAs which have capability continuity needs, the ECCEMP responds to the target capability calling request, and sends the instruction to ECP3 to send the result of target capability to UE associated ECA on ECP2.

3. The ECCEMP determines the ECP4 that has the target capability, based on - but not limited to - the location information of ECP2 and the status of target capability on ECP4.

NOTE 2 - The ECCEMP may consider the following information when determining the target capability: the location of the various ECPs and the ECA on ECP2; the load of the ECPs; the response delays of the ECPs.

4. The ECCEMP sends the capability calling instruction to ECP4 to send the capability calling result to the UE associated ECA on ECP2.

## **9. Considerations about charging of edge computing capability exposure**

The ECCEMP is responsible for the charging [b-ITU-T Y.2233] of capability exposure of edge computing capabilities and related authentication and authorization. The charging of capability exposure of edge computing capabilities takes into account the capability exposure usage information recorded by the ECP, including, but not limited to, the number of calls, the network traffic, the calling time and the priority information. Based on the usage information, metering information is generated, including but not limited to the capability caller identifier, the capability identifier, the calling times, the calling traffic amount.

## **10. Security considerations**

Concerning the capability calling and usage as specified in clause 8 and clause 9, the authentication and authorization of capability registration and capability calling are required. In addition, the ability to protect the capability information is required in order to avoid information leaking and unauthorized access.

The security and privacy related requirements specified in [ITU T Y.3101] [ITU-T Y.3105] are applicable to this Recommendation.

## **Bibliography**

- [b-ITU-T Y.2233] Recommendation ITU-T Y.2233 (2010), *Requirements and framework allowing accounting and charging capabilities in NGN*
- [b-ITU-T Y.3100] Recommendation ITU-T Y.3100 (2018), *Requirements of IMT-2020 network*
- [b-ITU-R M.1645] Recommendation ITU-R M.1645 (2003), *Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000.*
-