

Design for Edge/Fog Platform Management and Orchestration

Hung-Yu Wei¹, Yao Chiang, Tse-Yu Chen
Department of Electrical Engineering, National Taiwan University

1. Edge Platform Management and Orchestration²

The technology of edge computing has evolved considerably in recent years with the support of emerging key enablers like Network Function Virtualization (NFV) and Software Defined Network (SDN). NFV and SDN can be used to consolidate and deliver the networking components needed to support a fully virtualized infrastructure, such as virtual servers, storage units, and even other networks. They are applicable to any data plane processing or control plane function in both wired and wireless network infrastructures.

To solve the limited resource problem that is natural for edge computing systems, IEEE P1935 specifies the system architecture, resource management, application lifecycle operations, and communication specification between multiple Edge/Fog systems. The unified management and orchestration, such as the regulation of the transmissions of messages, the controlling and allocation of the resources of the edge devices, and the monitoring and arrangement in the mobile application lifecycle, can satisfy the mentioned needs and ensure better availability, flexibility, reliability, scalability, stability, service mobility, and performance. It means that the system shall be easy for multiple types of users to operate and configure, be hard to down while facing different challenges, and be as efficient as possible.

IEEE P1935 standard for Edge/Fog manageability and orchestration identifies the operations of resource management for Edge/Fog systems to enable more convenient and efficient deployment. As the core part of the evolving technologies of edge computing, the Edge platforms are responsible for handling various tasks, including computing, networking, and storage tasks. In order to process these tasks in a coordinated and scalable manner, the platforms need to communicate with other entities to share necessary information and collect needed resources. Aside from the Edge platforms, the Edge and Fog Orchestrator (EFO) and the Edge Platform Managers are also involved in the process. This document describes the RESTful service APIs of the resource operations which are used for the communication of these components, including resource creation, resource status query, resource discovery, resource reconfiguration, and resource deletion. For each operation, the document defines the procedures, conditions, involved information, and format of requests and responses, which can be found in the corresponding clause or subclause.

1.1 Overview of Edge Platform Resource Management

The service APIs are used for the resource management and communication between the components. As shown in Figure 3.1-1, these service APIs can be used to do several common operations on the resources, including resource creation, resource status query, resource search, resource reconfiguration, and resource deletion. The resources we mention here include physical resources and virtual resources. The physical resources are the hardware equipment available for the Edge/Fog system, including physical servers, facilities, and infrastructure. The virtual resources can be considered as the software, computing resources, networking resources, and storage resources in the system, as well as the virtualized infrastructure.

¹ Corresponding Author: Hung-Yu Wei (email: hywei@ntu.edu.tw)

² This document was submitted as technical contribution to IEEE P1935 working group in 2020. Part of the technical contents have been included in the IEEE P1935 standard.

There are totally five types of API supported:

- a) Resource creation: This operation API is used to create the target resource in the Edge/Fog system. It includes registration and configuration of both physical resources and virtual resources. The HTTP method POST is supported.
- b) Resource status query: This operation API is used to fetch the information of the certain resource. The HTTP method GET is supported.
- c) Resource discovery: This operation API is used to search the resources that meet the specific filtering rules. The HTTP method GET is supported.
- d) Resource reconfiguration: This operation API is used to modify the properties and attributes of the target resource after its creation. There are two types of HTTP methods supported, PUT and PATCH, which are used to replace the whole information and to update partial information respectively.
- e) Resource deletion: This operation API is used to remove the target resource in the Edge/Fog system. The HTTP method DELETE is supported.

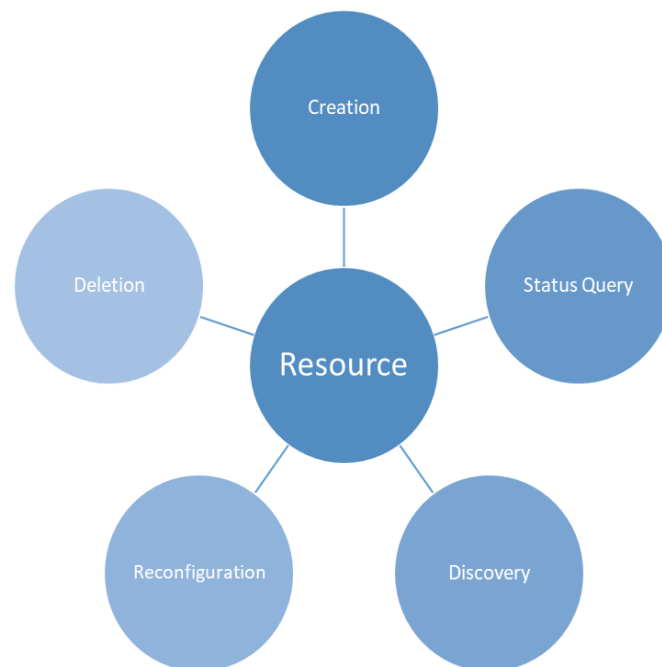


Figure 1—The common operations of the Edge system resource management

These service APIs for the Edge/Fog platforms are RESTful HTTP APIs, which should be independent of technology implementation. All the resources in the Edge/Fog system have their URLs to be called by the APIs. The specification of service APIs should include the following information:

- The purpose of the API: the description and usage of the operation API.
- Request format: the format of the request. The format includes the HTTP methods, headers, and payload.
- Response format: the format of the response, including the supported HTTP status codes.
- HTTP methods supported: the HTTP method(s) used in the operation API. An operation API may support multiple HTTP methods.
- Representation supported: most of the operation API payload should be in JSON format. The exceptions should be specified if they exist.

The components that are mentioned in Section 3.2 are listed in Table 3.1-1. The operation APIs can be sent from the internal VF M&O. It may also be possible for some external API clients which are able to send HTTP requests to send the requests, but additional verification and authorization are needed in that case. In this section, we only discuss the situation where the request is sent from the VF M&O.

Table 1—The components mentioned in 1.2

Types	Components	1.2.1	1.2.2	1.2.3	1.2.4	1.2.5
Users	Service Provider					
	Operator					
	End User App					
EFO	External API					
	End User Proxy					
	Edge App Designer					
	VF M&O	V	V	V	V	V
	Edge Inventory Manager					
	LCM Enabler					
	Rule Framework					
Control nodes	EFO Controllers					
	Edge Platform Manager					
Compute nodes	IM/VIM	V	V	V	V	V
	Edge Platform					
	Edge App					
	VNI	V				V
	Data Plane					

Table 2—The terms used in the operation APIs

Terms	Descriptions	Clauses
Attribute	Attributes are the additional information of the resource.	1.2.4
Edge resource	Edge resources are the resources created in the Edge system. These resource can be further sliced as the resource quota.	1.2.1
Filtering rule	Filtering rules are the restrictions and conditions to narrow down the search results.	1.2.3
Physical resource	Physical resources are the physical hardware of the Edge system, for example, the physical servers.	1.1
Property	Property means the characteristics of the resource.	1.2.4
Resource	Resources are the physical and virtual assets that can be leveraged by the Edge system.	1.1
Resource configuration	The resource configuration is a part at the end of the resource creation operation.	1.2.1
Resource creation	It is the operation used to create new resources in the Edge system.	1.2.1
Resource deletion	It is the operation used to delete resources in the Edge system.	1.2.5
Resource discovery	It is the operation used to search and return the list of resources matching the filtering rules.	1.2.3
Resource quota	After resource creation, the Edge resources will be created and configured. In the configuration process or the resource reconfiguration operation later, the Edge resource can be sliced into resource quota.	1.2.1
Resource reconfiguration	It is the operation used to modify the information of resources. There are two types of this operation: resource replacement and resource update.	1.2.4
Resource replacement	It is one of the two types of modification operation which replaces whole information of the resource.	1.2.4
Resource status query	It is the operation used to get the current information of the target resource.	1.2.2
Resource update	It is one of the two types of modification operation which only updates partial information of the resource.	1.2.4
Resource URL	Resource URLs are the unique addresses of the resources for the Edge system to locate.	1.1
Virtual resource	Virtual resources are the non-physical resources in the Edge system, including computing resources, networking resources, storage resources, etc.	1.1

1.2 Resource Manageability Procedures

This section describes the service APIs which can be used to do several resource operations, including resource creation, resource status query, resource discovery, resource reconfiguration, and resource deletion. The request procedures start from the VF M&O in the EFO, and is processed by the VIM in the most cases. In this section, we will describe the usage, procedures, as well as the format of the requests and responses of each operation.

1.2.1 Resource Creation

New resources can be created by sending an HTTP POST request to the system. The POST request should contain the information about the new resource that will be created. The new resources called “Edge/Fog resources” will be given a unique identifier (ID) named the resource URL. After creation, these Edge/Fog resources will be configured automatically based on the request payload as well. The resource may be split into several pieces or slices, called as “resource quota”, providing higher flexibility

and system utilization. For example, the virtual resources contained in one physical server can be cut into slices to be leveraged in different fields, and these slices are called as the resource quota of the physical resource (that physical server). The other details of the configuration process vary from system to system, and are therefore beyond the scope of this standard.

Such resource creation operations can be made for both physical and virtual resources. The resource creation operation registers the real information of the physical resources to the Edge/Fog system to make them visible and usable for the system, and connect the virtual resources to the physical resources which host them. Such connection is shown in a way where the virtual resources are the child resources of the physical resources as the parent resources. The connections can be modified by the resource reconfiguration operation.

VF M&O seldom makes the resource creation operation on its own, but it is possible for it to create and connect the virtual resources to the physical resources automatically after the creation of the physical resources.

The pre-condition of the procedure:

- The target resource is not yet created in the Edge system.

The post-condition of the procedure:

- The resource has been created and configured as network infrastructure in the Edge system.

The procedures of the operation:

- a) The VF M&O sends a resource creation request to the IM/VIM (Table 3).
- b) The IM/VIM computes the resource URL and then creates the resource as the Network Infrastructure. The IM/VIM configures the properties and attributes of the resource based on the payload of the request, the process of which differs from system to system.
- c) After the resource is created either successfully or failingly, the IM/VIM sends the response to the VF M&O (Table 4 & Table 16).

If the server creates the new resource successfully, the HTTP response status “201 created” shall be returned, which includes the path information of the created resource (Table 4). On failure, the specific status error code shall be returned (Table 16).

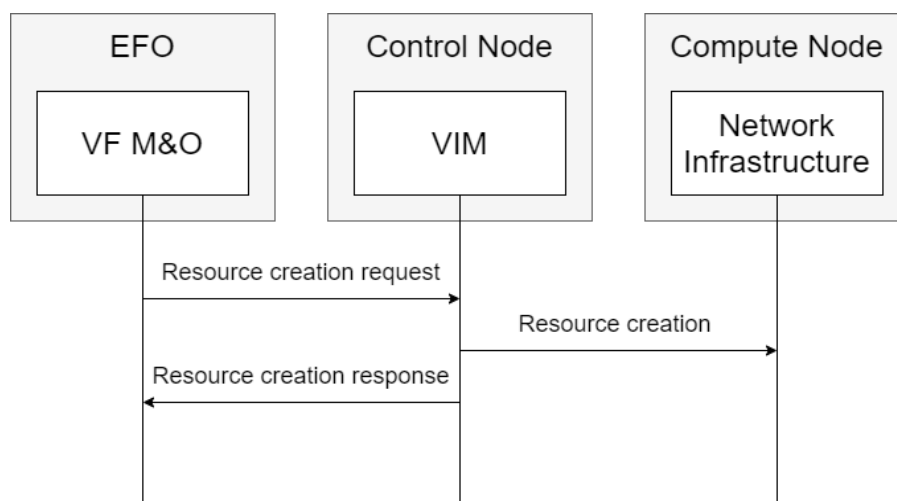


Figure 2—The workflow of the resource creation

Table 3—Resource creation request

Element	Description
Request Type	HTTP POST
Parent Resource ID	The identifier of the parent resource of the created resource. It shall be contained in the URL.
Request Payload	The payload contains the information and data of the created resource.

Table 4—Resource creation successful response

Element	Description
HTTP Status Code	201 Created
HTTP Header	URL The URL of the created resource.
Response Payload	(Empty)

1.2.2 Resource Status Query

The resource information can be obtained via an HTTP GET request that ought to be available for most resources and contains the empty payload. This operation is usually used when the entity already knows the URL of the target resource, which can be obtained via the resource discovery operation described in Section 3.3. The resource status query operation can be used for both physical and virtual resources. In all cases, this operation only returns the information of the target resource, and it will give corresponding error messages if the requested resource is not found in the system. This operation is also the necessary step before the resource reconfiguration operation described in Section 3.4 to avoid race conditions.

The pre-condition of the procedure:

- The target resource has been created in the Edge system.

The post-condition of the procedure:

- The entity which sent the request gets the information of the target resource.

The procedures of the operation:

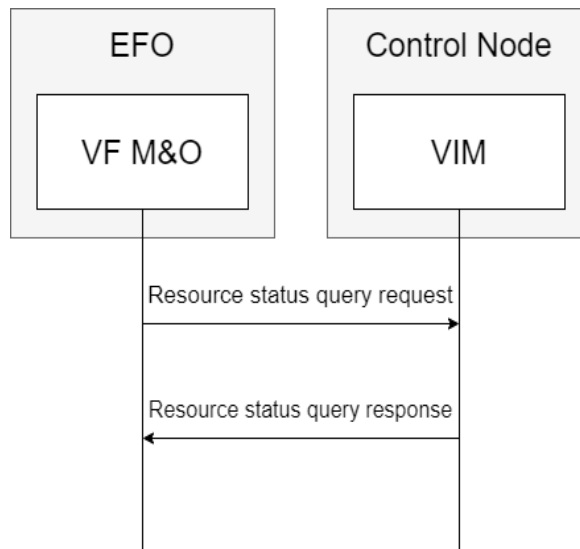


Figure 3—The workflow of the resource status query

- The VF M&O sends the resource status query request to the IM/VIM (Table 5) in order to query the target resource.

- b) If the specific resource is found, the information of the requested resource shall be returned to the VF M&O as an HTTP response (Table 6). On failure, the specific error status code shall be returned (Table 16).

On success, the “200 OK” status code shall be returned with information of the requested resource (Table 6); on failure, the specific error status code shall be returned (Table 16).

Table 5—Resource status query request

Element	Description
Request Type	HTTP GET
Resource ID	The identifier of the <u>target</u> resource. It shall be contained in the URL.
Request Payload	(Empty)

Table 6—Resource status query successful response

Element	Description
HTTP Status Code	200 OK
HTTP Header	(Empty)
Response Payload	The information of the requested resource, including the version number of it.

1.2.3 Resource Discovery

The resource searching with the filtering rules can be done by an HTTP GET method with query parameters that can be used to control the content of the query result. The filtering rules are the restrictions and conditions to narrow down the search results. The GET request is with the empty payload. Compared to the resource status query operation, the resource discovery operation is able to return all the resources that match the filtering rules. If there is no resource that matches the filtering rules, an empty list should be returned.

As the other operations, the resource discovery operation is suitable for both physical and virtual resources. The users and the system can use this operation to find all the resources they need via the VF M&O respectively.

The pre-condition of the procedures:

- The target resource has been created in the Edge system.

The post-condition of the procedures:

- The entity which sent the request receives all the resources matching the filtering rules as well as their information.

The procedures of the operation:

- a) The VF M&O sends a resource discovery request to the IM/VIM (Table 7).
- b) The IM/VIM searches and filters the resources to return a list of resources which satisfy the given filtering rules as an HTTP response to the VF M&O (Table 8). On failure, the specific error status code shall be returned (Table 16).

On success, the “200 OK” status code shall be returned with the payload that contains the specific resource data, adjusted according to the passed parameters (Table 8). On failure, the specific error status code shall be returned (Table 16). If no resource matches the filtering rules, it is considered as the failure case.

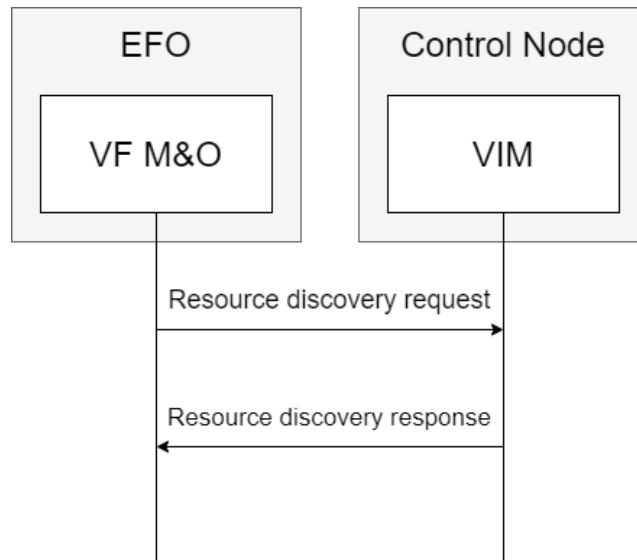


Figure 4—The workflow of the resource discovery

Table 7—Resource discovery request

Element	Description
Request Type	HTTP GET
Filtering Rules	The restrictions and conditions to narrow down the search results. They shall be contained in the URL.
Request Payload	(Empty)

Table 8—Resource discovery successful response

Element	Description
HTTP Status Code	200 OK
HTTP Header	(Empty)
Response Payload	A list of the resources that match the filtering rules. It shall be in JSON format.

1.2.4 Resource Reconfiguration

Resource reconfiguration means the modification of the properties or attributes of the target resource. These properties and attributes are generated during the resource creation operation automatically, based on the payload of the resource creation operation request. There are two methods to reconfigure a resource: replacement and update. For ease of differentiation, these two methods can be referred to as “resource replacement” and “resource update” respectively. The resource replacement operation replaces the whole information of the target resource, and the resource update operation only updates the information that is contained in the request payload.

For the resource replacement, the resources can be overwritten by the HTTP PUT method with a resource representation. That is, all the old resource content will be omitted after resource replacement. For the resource update, the HTTP PATCH method is used to update a resource by only changing the part described in the payload from the client, leaving the other part unchanged. PATCH requests are available for all situations where PUT requests are also available. In both cases, before the modification, the resource status query operation should be performed to ensure the version of the resource, in order to avoid race conditions.

As the other operations, the resource discovery operation is suitable for both physical and virtual resources. For example, the Edge Service Operators may need to either update the information of their physical servers, change the settings of the computing resources, or re-slice the virtual resources. The VF M&O can also be set up to update the information automatically based on the feedback from the system.

1.2.4.1 Conditions and Involved Information

The pre-condition of the procedure:

- The target resource has been created in the Edge system.

The post-condition of the procedure:

- The properties and attributes of the target resource will be modified. If needed, the related resources will be reconfigured as well.

The procedures of the resource replacement operation:

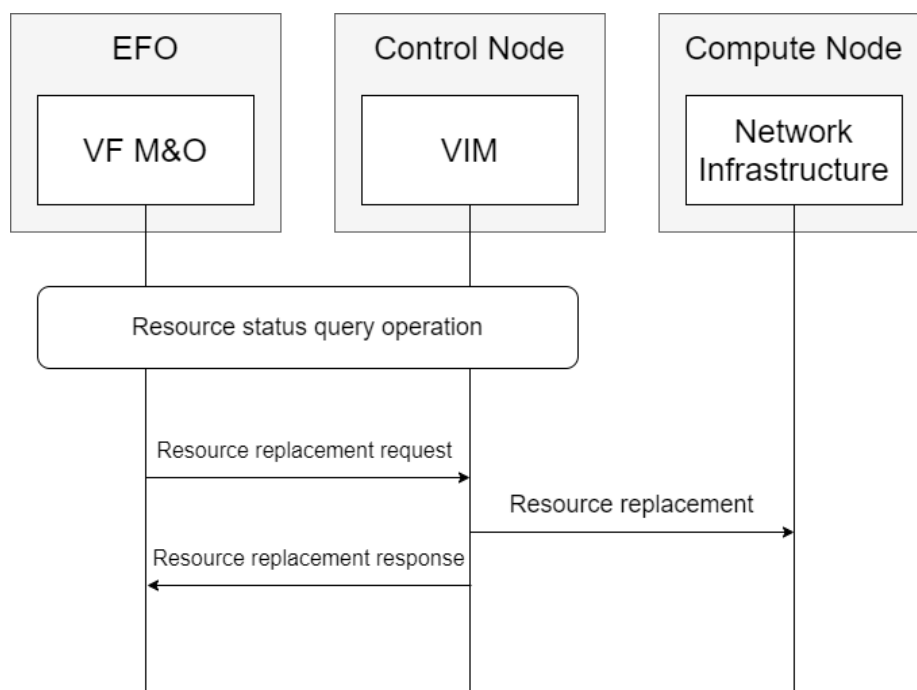


Figure 5—The workflow of the resource replacement

- Before the resource replacement, the VF M&O needs to make the resource status query operation to get the version number of the target resource as described in 1.2.2.2. If the specific resource is found, the information of the requested resource shall be returned. If not, the process is terminated as there is no resource to replace.
- The VF M&O sends the resource replacement request (Table 9) which contains the version number obtained in Step 1 to the IM/VIM.
- The IM/VIM checks the version number of the resource. If the version number is valid, the IM/VIM will replace the target resource. If not, an error message will be returned. If needed, the related resources will be reconfigured as well.
- The IM/VIM returns the success or error message to the VF M&O (Table 11, Table 12, or Table 16).

The procedures of the resource update operation:

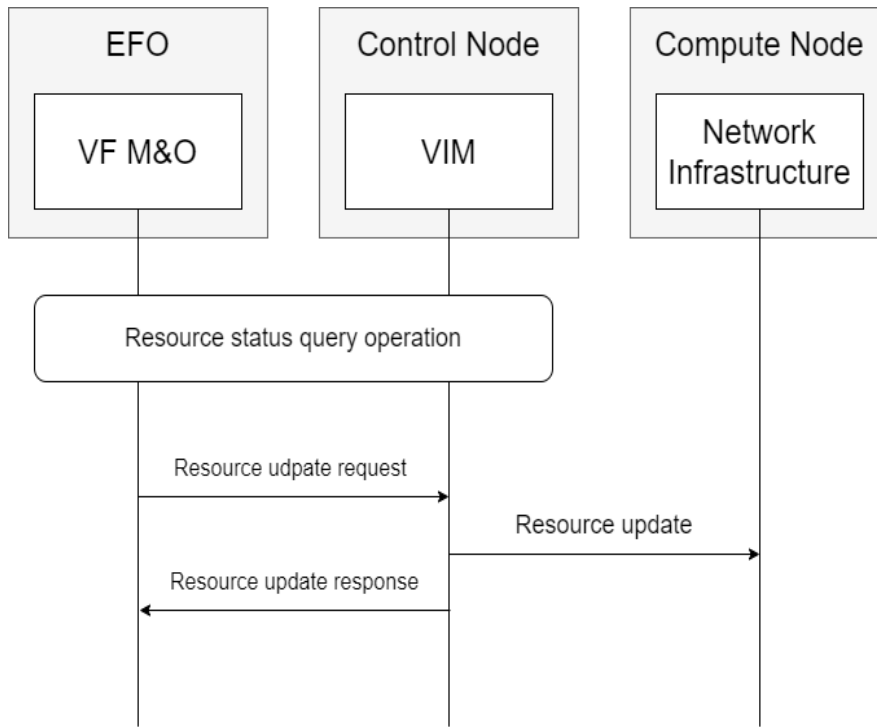


Figure 6—The workflow of the resource update

The request can be sent from the API client or the VF M&O.

- a) Before the resource replacement, the VF M&O needs to make the resource status query operation to get the version number of the target resource as described in 1.2.2.2. If the specific resource is found, the information of the requested resource shall be returned. If not, the process is terminated as there is no resource to replace.
- b) The VF M&O sends the resource update request (Table 10) which contains the version number obtained in Step 1 to the IM/VIM.
- c) The VIM checks the version number of the resource. If the version number is valid, the IM/VIM will update the target resource. If not, an error message will be returned. If needed, the related resources will be reconfigured as well.
- d) The IM/VIM returns the success or error message to the VF M&O (Table 11, Table 12, or Table 16).

On success, either “200 OK” or “204 No Content” shall be returned (Table 11 & Table 12); if a client attempts to modify the resource with an expired version number, “412 Precondition Failed” shall be returned. On other failures, the specific status error code shall be returned (Table 16).

Table 9—Resource replacement request

Element	Description
Request Type	HTTP PUT
Request Payload	The new data and information of the target resource. The all old content of the target resource will be replaced by the content of this payload.
Version Number	The version number of the target resource. It is used to avoid the race conditions.

Table 10—Resource update request

Element	Description
Request Type	HTTP PATCH
Request Payload	The new data and information of the target resource. Only the resource attributes that are included in the payload will be updated, and all other attributes will remain unchanged.
Version Number	The version number of the target resource. It is used to avoid the race conditions.

Table 11—Resource reconfiguration successful response with non-empty payload

Element	Description
HTTP Status Code	200 OK
HTTP Header	(Empty)
Response Payload	The new information of the updated resource.

Table 12—Resource reconfiguration successful response with empty payload

Element	Description
HTTP Status Code	204 No Content
HTTP Header	(Empty)
Response Payload	(Empty)

1.2.5 Resource Deletion

The HTTP DELETE method is used to delete the requested resource. The request and the response are both with an empty payload, but it is also possible for the response to return the final information of the resource. After deletion, all the information about the deleted resource will be removed, and the related resources will be updated and rearranged.

As the other operations, the resource discovery operation is suitable for both physical and virtual resources. The system can delete the target resource based on their needs and the feedback. The deletion of physical resources means that these resources are removed from and no longer visible to the system, but the physical servers may be still at the same location. The deletion of virtual resources removes the resource actually; however, it is possible to re-create them as long as the physical resource is still in the Edge/Fog system.

1.2.5.1 Conditions and Involved Information

The pre-condition of the procedure:

- The target resource has been created in the Edge system.

The post-condition of the procedure:

- The target resource will be removed from the Edge system. The related resources are reconfigured.

The procedures of the resource deletion operation:

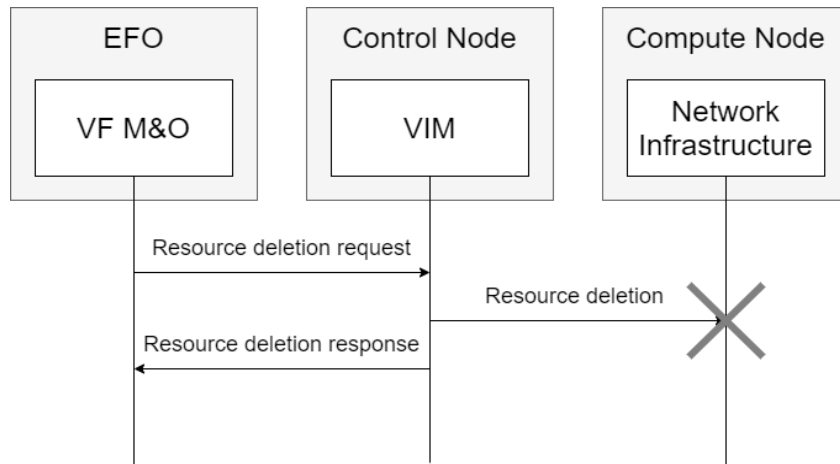


Figure 7—The workflow of the resource deletion

- a) The VF M&O sends a resource deletion request to the IM/VIM (Table 13).
- b) The IM/VIM deletes the target resource as the Network Infrastructure, and clean up the related information. If the resource is the parent of other resources, it is optional to delete those resources as well. The related resources will be reconfigured.
- c) The IM/VIM returns the deletion success or error message to the VF M&O (Table 14, Table 15, or Table 16).

On success, if the payload is empty, “204 No Content” shall be returned, otherwise “200 OK” shall be returned (Table 14 & Table 15). On failure, the specific error code shall be returned (Table 16).

Table 13—Resource deletion request

Element	Description
Request Type	HTTP DELETE
Request Payload	(Empty)

Table 14—Resource deletion successful response with non-empty payload

Element	Description
HTTP Status Code	200 OK
HTTP Header	(Empty)
Response Payload	The information of the deleted resource.

Table 15—Resource deletion successful response with empty payload

Element	Description
HTTP Status Code	204 No Content
HTTP Header	(Empty)
Response Payload	(Empty)

1.3 Summary

The issues of resource management and allocation has always been a major focus of research in areas like system designs. Different to the traditional cloud computing systems and platforms, however, Edge/Fog systems are expected to be deployed in a distributed approach widely and thus have fewer resources. The characteristic of limited resources of Edge/Fog systems makes the resource management issues further vital.

This document describes the procedures of resource management in the IEEE P1935 Edge/Fog systems, including not only the necessary files involved in such procedures but also the specification

for relevant requests and responses. It explains the workflows of the five main resource management procedures, resource creation, resource status query, resource discovery, resource reconfiguration, and resource deletion; and the details can be found in corresponding subclauses.

The detailed information can help practitioners to implement a unified, operational Edge/Fog system without the need to start from scratch. The uniform specifications introduced in the document should help the regulation and integration of Edge/Fog systems and further improve the development of related industries.