Draft new Recommendation ITU-T Y.3120 (ex.Y.IMT2020-fa-lg-lsn)

Functional Architecture for latency guarantee in large scale networks including IMT-2020 and beyond

Summary

This Recommendation specifies the functional architecture, functional entities, reference points, and operational procedures, for the requirements and framework defined in Y.3113, based on the architecture defined in Y.2111. Meanwhile, Y.3113 specifies the use of flow aggregate (FA)-based scheduling and regulators at aggregation domain (AD) boundaries. Y.2111 specifies the resource and admission control functions (RACF) in support of end-to-end quality of service (QoS) and necessary transport functions in next generation networks (NGNs).

Keywords

Latency guarantee, large scale network, flow aggregate, quality of service, regulator

Table of Contents

Draft new Recommendation ITU-T Y.3120 (ex.Y.IMT2020-fa-lg-lsn)

Functional Architecture for latency guarantee in large scale networks including IMT-2020 and beyond

1 Scope

This Recommendation specifies the architecture and procedures for latency guarantee in large scale networks, based on the requirements and framework specified in ITU-T Y.3113, as follows:

- Functional architecture
- Functional entities and reference points
- Operational procedures

Detailed protocols for the reference points are out of scope of this Recommendation.

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-R M.1645] Recommendation ITU-R M.1645 (06/2003), *Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000*.

[ITU-R M.2083] Recommendation ITU-R M.2083-0 (09/2015), *IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond*.

[ITU-T E.800] Recommendation ITU-T E.800 (09/2008), *Definitions of terms related to quality of service*.

[ITU-T Y.2111] Recommendation ITU-T Y.2111 (2011), *Resource and admission control functions in Next Generation Networks.*

[ITU-T Y.3113] Recommendation ITU-T Y.3113 (2021), *Requirements and framework for latency guarantee in large scale networks including IMT-2020 network.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 IMT-2020 [ITU-R M.2083]: Systems, system components, and related technologies that provide far more enhanced capabilities than those described in [ITU-R M.1645].

NOTE – [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 customer premises equipment [ITU-T E.800]**:** Telecommunications equipment located at the customer installation on the customer side of the network interface.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 aggregation domain: A maximal set of the interfaces of the consecutive relay nodes in the path, travelled by a flow, in which the 'flow membership' of the flow aggregate the flow belongs to is unaltered.

NOTE – An aggregation domain is defined per flow.

3.2.2 domain: A set of relay nodes and end-hosts under a single administrative control or within a closed group of administrative control; these include campus wide networks, private WANs, and IMT-2020 networks.

NOTE – This definition references the description in Introduction clause of [b-IETF RFC 8655].

3.2.3 large scale network: A network or a set of interconnected networks, with diameter of 16 or larger, in which the numbers of flows and nodes are proportional to the diameter of the network.

3.2.4 relay node: A node supporting relay functionality that acts as an intermediary node, through which other nodes can pass their traffic (e.g. router, switch, gateway, etc.).

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AD	Aggregation Domain
AN	Access Network
ATS	Asynchronous Traffic Shaping
CPE	Customer Premises Equipment
CN	Core Network
DiffServ	Differentiated Services
E2E	End-To-End
FA	Flow Aggregate
FIFO	First-In First-Out
IntServ	Integrated Services
IR	Interleaved Regulator
PD-FE	Policy Decision Functional Entity
PE-FE	policy Enforcement Functional Entity
PFAR	Port-based Flow Aggregate Regulator
QoS	Quality of Service
RACF	Resource and Admission Control Function
RSpec	Reserve Specification
RSVP	Resource reSerVation Protocol

SCF	Service Control Functions
TRC-FE	Transport Resource Control Functional Entity
TRE-FE	Transport Resource Enforcement Functional Entity
TSN	Time Sensitive Network
TSpec	Traffic Specification

5 Conventions

The keywords “is required to” indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The keywords “is recommended” indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

The keywords “can optionally” indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor’s implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6 Introduction

Latency sensitive applications across multi-domain large scale networks emerge, such as augmented reality, virtual reality, tactile internet, and smart industry. [ITU-T Y.3113] describes the requirements and framework for latency guarantee in large scale networks. Y.3113 specifies the framework for latency guarantee in such networks, with the combination of the flow aggregate (FA)-based queuing and scheduling architecture and the regulators at the aggregation domain (AD) boundaries.

At the network configuration phase, the boundaries of ADs should be decided. The size of an AD is a key network configuration parameter. It affects the number of FAs, number of flows in an FA, the number of regulators, and the end-to-end (E2E) latency bound itself. In the call setup phase, given the traffic specification of a flow, the E2E latency bound must be pre-calculated with the cooperation among networks. An FA may have flows join/leave dynamically, therefore it is necessary to re-negotiate the E2E latency bounds with the sources of flows in the FA.

[ITU-T Y.2111] specifies the resource and admission control functions (RACF) in support of end-to-end quality of service (QoS) and necessary transport functions in next generation networks (NGNs). The functional architecture of this Recommendation enhances that of Y.2111 for latency guarantee in large scale networks.

This Recommendation specifies the functional architecture, functional entities, reference points, and operational procedures for the framework defined in Y.3113, based on the functional architecture defined in Y.2111.

7 Functional architecture

Figure 1 depicts the functional architecture specified in this Recommendation, based on Y.2111. This Recommendation extends functional entities and reference points defined in Y.2111, such that they can guarantee latency upper bounds for requesting flows in large scale networks possibly over multiple administrative domains.

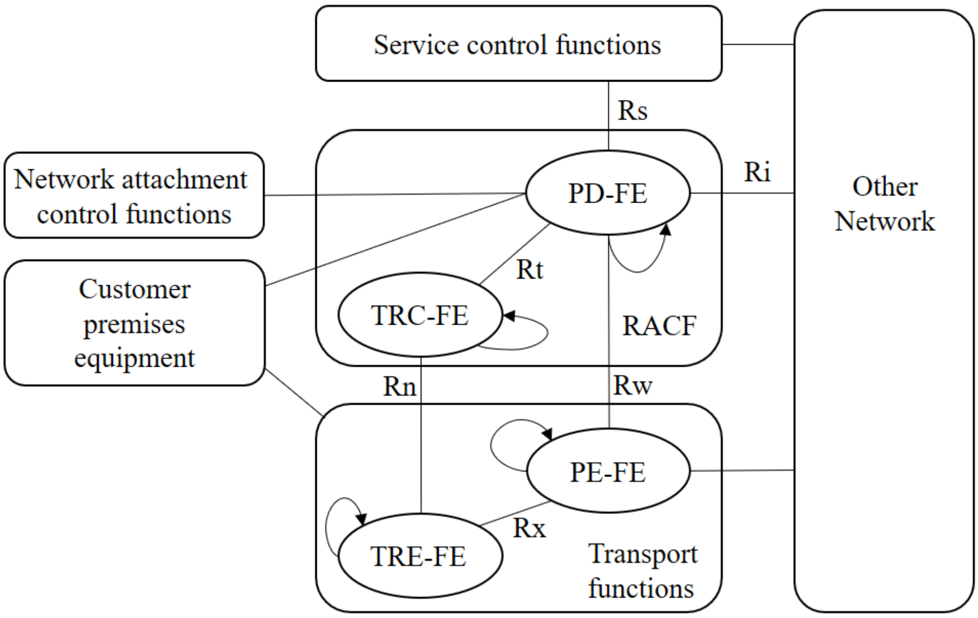

Figure 1 – Functional architecture for latency guarantee

The RACF acts as the arbitrator between service control functions and transport functions for QoS related transport resource control [ITU-T Y.2111]. The policy decisions made by the RACF are based on transport subscription information, service level agreements (SLAs), network policy rules, service priority, and transport resource status and utilization information.

The architecture in Figure 1 includes SCF (service control functions), PD-FE (policy decision functional entity), TRC-FE (transport resource control functional entity), TRE-FE (transport resource enforcement functional entity), PE-FE (policy enforcement functional entity), and NACF (network attachment control functions).

The PD-FE is required to provide a single point of contact to the SCF over Rs. For scalability in larger domains, multiple instances of PD-FE may be deployed, each one handling a subset of the PE-FEs. As a result, the PD-FE instance that receives a request over the Rs reference point may not be able to directly reach the PE-FE concerned. Hence the instances of PD-FE need to inter-communicate. It also allows the interaction between the TRC-FE instances within the same administrative domain. A single RACF can similarly interact with a single or multiple pair of PE-FE and TRE-FE for scalability. The details of the functional entities and reference points can be found in Y.2111.

The AD is the key concept for designing the network architecture for latency guarantee, as it is defined in Y.3113. Within an AD, flows having the same path are considered undistinguishable and are treated as a single FA. At the boundary of an AD, the flows are required to be regulated according to their traffic specifications (TSpecs).

An AD can encompass a single or multiple switching nodes. In an extreme case, an AD includes only a single hop, i.e. an output port of a node and the regulator at the next node. In this case an AD can be covered by a single RACF and a TRE-FE/PE-FE pair. In another extreme case, an AD includes an administrative domain. In this case an AD can be covered by multiple RACF instances, if the domain employs multiple RACFs for scalability. If an AD is covered by multiple TRE-FE instances, then the FA resource control parameters such as aggregated rates and burst size should be exchanged among the TRE-FE instances. The RACF and transport functions should be able to interact with other networks such that E2E latency of flows can be bounded. The RACF should be able to control the overall latency bounds of flows according to their RSpecs in a network.

The operational procedure in this Recommendation is composed of the three sub-operations; network configuration, call setup, and flow treatment at the data-plane. At the network configuration phase, the ADs should be decided. The size of an AD is a key network configuration parameter. It affects the number of FAs, number of flows in an FA, the number of regulators, and the E2E latency bound itself. FA granularity and membership should also be decided. In the call setup phase, given the TSpec of a flow, the E2E latency bound must be calculated with the cooperation among ADs before the admission. An FA may have flows join/leave dynamically, therefore it can be necessary to re-negotiate the E2E latency bounds with the sources of flows in the FA.

8 Functional entities and reference points

The functional architecture, functional entities, and reference points described in this Recommendation are based on Y.2111. In this Clause, it is described how it should be extended to provide latency bounds guarantee for requesting flows.

8.1 Functional entities

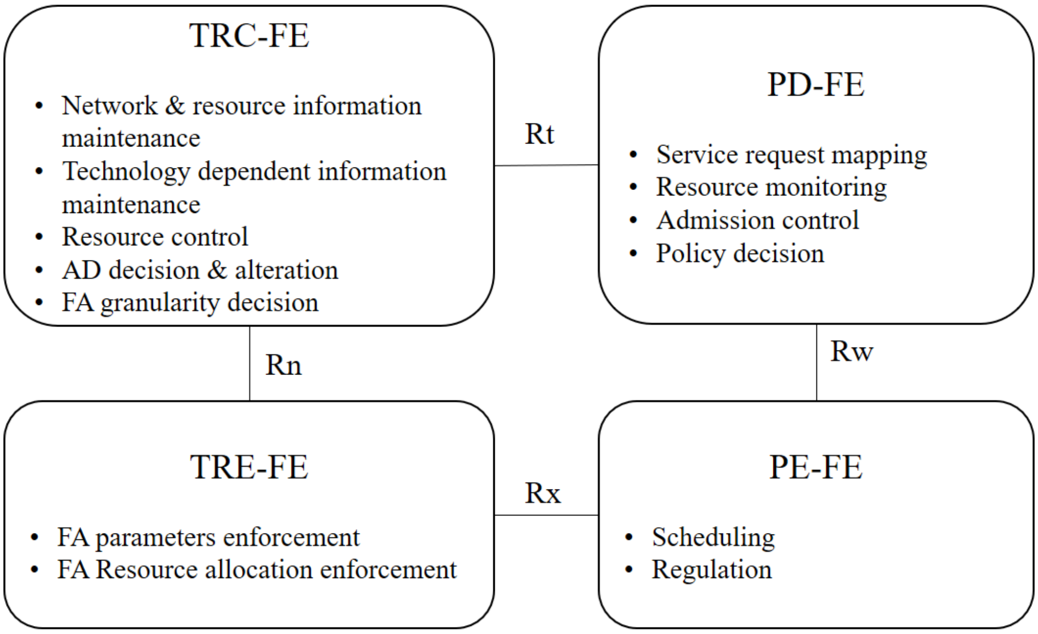

Figure 2 – Functional entities and their roles for latency bound guarantee

Figure 2 depicts the four major functional entities and their roles, and the reference points between the entities, for latency bound guarantee.

- The PD-FE handles the QoS resource requests received from the SCF via the Rs reference point or from the PE-FE via the Rw reference point. The PD-FE decides the admission of a flow. It monitors the available resources, maps the service request to the network resource, and decides the policy regarding the resource allocation to the flows.
- The TRC-FE collects and maintains the network topology information and resource status information. It is also responsible for the AD boundary decision, FA membership decision, and related information maintenance. The TRC-FE should be able to perform AD boundaries decision and alteration. It should be able to decide the granularity and the resource allocated to the FA.
- The PE-FE in the transport functions enforces the network policy rules instructed by the PD-FE on a per-subscriber, per-flow, or per-FA basis. The PE-FE includes functions such as regulation, service rate allocation, scheduling, packet filtering, traffic classification and marking, traffic policing, as well as collecting and reporting resource usage information.
- The TRE-FE enforces the transport resource policy rules instructed by the TRC-FE at the FA level with proper link layer protocols (e.g., virtual LAN, virtual private network, and multi-

protocol label switching). It should be able to perform the resource enforcement functions based on link layer information. The TRE-FE is recommended to dynamically decide/modify the TSpec e.g. service rate associated with an FA as its membership changes, and to set traffic management parameters such as average rate or burst size of the data link layer protocols. The amount of the bandwidths allocated to FAs should be controlled such that the latency bounds in an AD for the FAs are adjustable.

A single set of the TRE-FE and PE-FE can be allocated to a single AD, or multiple ADs. If the network has the capability of merging and dividing ADs as described in Clause 9.1.3, then a single TRE-FE and PE-FE pair should be able to handle multiple ADs. Through the Rx the parameters for scheduling and regulation are delivered from the TRE-FE to the PE-FE.

8.2 Reference points

The reference points for latency guarantee, referred in Figure 1, are described in the following.

- The Rs reference point allows resource request information needed for resource authorization and reservation, for latency bound guarantee, to be exchanged between the PD-FE and the SCF. The Rs reference point provides the ability for the SCF to make requests for resource authorization/reservation and QoS/priority handling of flows. The Rs reference point may operate as an intra-domain or an inter-domain reference point.
- The Ri reference point conveys the information on QoS handling, priority handling, resource usage, AD boundaries, and FA granularities between network administrative domains. The Ri reference point is an inter-domain reference point.
- The Rw reference point allows the final admission decisions made by the PD-FE to be realized with appropriate actions at the PE-FE. The Rw reference point allows the PD-FE to push the admission decisions to the PE-FE to provide latency guarantee, and also allows the PE-FE to request the admission decisions when path-coupled resource reservation mechanisms are in use. The PD-FE may specify resources to be reserved and/or committed for flows and QoS handling such as packet marking and regulation to use. The Rw reference point is an intra-domain reference point.
- The Rn reference point allows the TRC-FE to collect the network topology and resource status information of an access or a core network. The TRC-FE makes the decision on the AD boundaries based on the information collected. The Rn reference point is an intra-domain reference point.
- The Rt reference point allows the PD-FE to interact with the TRC-FE to detect and determine the requested QoS resource of flows along the flow path in the involved access network and core network. The TRC-FE collects and maintains the network topology information and resource status information, and sends it to PD-FE through Rt. The PD-FE may also request that the TRC-FE provides the path selection and flow aggregation information for a given flow in the core network. The Rt reference point is an intra-domain reference point.
- The Rx reference point allows the PE-FE to collect enforcement parameters for FAs from TRE-FE. The information collected by the Rx reference point and the policies by the Rw reference points combined will render the precise resource control for flows. The Rx reference point is an intra-domain reference point.

9 Operational procedures

The procedure for latency guarantee in large scale networks includes three sub-procedures; network configuration, call setup, and flow treatment at the data-plane.

In the network configuration procedure, aggregation domains and flow aggregation granularities are determined. The network topology, network size, number of flows, and the capability of nodes are taken into consideration. The size of an AD is a key network configuration parameter. It affects the number of FAs, number of flows in an FA, the number of regulators, and the E2E latency bound itself. The network configuration procedure does not alter the network topology.

The call setup procedure in this Recommendation is based mainly on ITU-T Y.2111. In the call setup phase, given the TSpec of a flow, the E2E latency bound must be pre-calculated with the cooperation among network administrative domains, if necessary. An FA may have flows join/leave dynamically, therefore it is necessary to re-negotiate the E2E latency bounds with the sources of flows in the FA.

The flow treatment procedure, which includes queuing, scheduling, regulation, etc., in this Recommendation is based on ITU-T Y.3113. The queuing and scheduling are per-FA basis. The first-in first-out (FIFO) characteristic of an FA within an AD shall be kept, in order for the subsequent interleaved regulator (IR) not to increase the worst latency.

9.1 Network configuration procedure

Network configuration procedure determines the AD boundaries and the FA granularities, given the network topology. For this task, the information regarding network topology, relay nodes capabilities, and overall resource usage on each link have to be gathered by TRC-FE through reference points Rn. The RSpecs of admitted flows also have to be gathered through Rt. The TRC-FE is the final decision point of the AD boundaries and FA granularities.

Figure 3 describes the overall procedure for the network configuration.

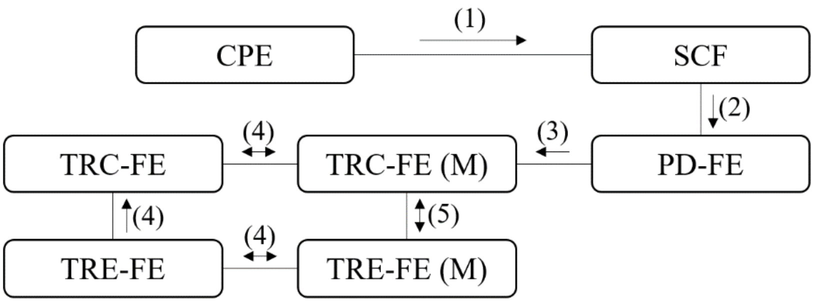

Figure 3 – Overall network configuration procedure

1) CPE requests application-specific service to SCF with optionally a service QoS requirements.

2) SCF identifies flows corresponding to the CPE request, and extracts or derives the RSpec of the flows, then requests to PD-FE for the network configuration and admission control.

3) PD-FE requests the master TRC-FE for the network configuration. Optionally TRC-FE can initiate the network configuration procedure when detecting network status changes.

4) The master TRC-FE gathers information regarding the network topology, capabilities and resource usage status, from other TRE-FEs through TRC-FE instances on the E2E path of the flows. This process can also be done through the interfaces between TRE-FEs.

NOTE: The master functional entity, denoted as FE(M), is the one that is responsible for the task specific to a flow. It requests the peer functional entity instances for the task.

5) TRC-FE also gathers, from the co-located master TRE-FE, information regarding the network topology, capabilities, and resource usage status. After the call setup procedure is completed by PD-FE, the master TRC-FE finally decides the ADs and FAs for the flows.

9.1.1 Aggregation domain (AD)

An AD is defined per a flow. An AD is a maximal set of the interfaces of the consecutive relay nodes in the path, travelled by a flow, in which the ‘flow membership’ of the flow aggregate the flow belongs to is unaltered. There should be one or more non-overlapping ADs in an end-to-end path of a flow. Requirement 7 in ITU-T Y.3113 specifies that it is required that networks be able to handle FAs as control elements. This requirement mandates that the network relay nodes should be able to queue and schedule a packet according to FA. An important consequence of such treatment based on FA is that the FIFO characteristic of the packets within an FA is maintained. In an extreme case in which packets are treated only based on their class, such as in DiffServ, an AD covers only a single node. In this case, the membership of an FA is unaltered only for a single hop. Note that this case still meets Requirement 7 in ITU-T Y.3113.

Flows with an identical path may have different ADs. An AD should have a regulation function at its segregation point. Figure 4 depicts a scenario in which the flows with the same path (flow 1 and 2) have different ADs thus be put into different FAs.

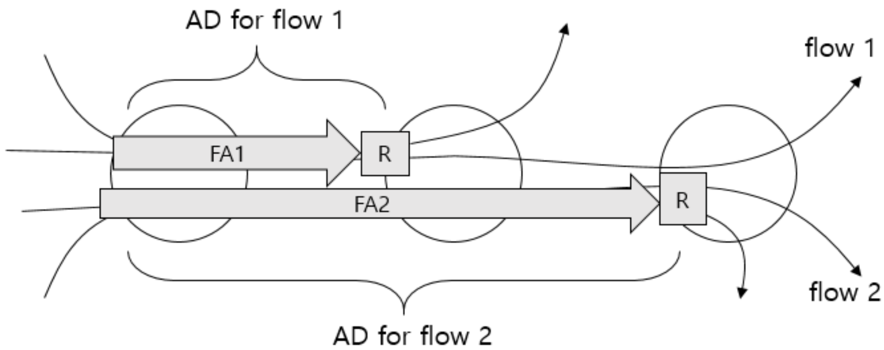

Figure 4 – **Aggregation domain for each flow with regulation function at the segregation point**

For operational efficiency, however, the flows with the same path and similar TSpecs are recommended to belong to a single FA and have an identical AD.

An aggregation point of an AD is defined to be a functional entity, at which the flow is aggregated into an FA. An aggregation point is defined per flow. An exemplary location of an aggregation point is an output port of a relay node. An aggregation point is part of an AD.

A segregation point of an AD is defined to be a functional entity, at which the FA is segregated. The flows are separated into different output paths. A segregation point is defined per flow. An exemplary location for a segregation point is an input port (or a switch fabric) of a relay node. A segregation point is part of an AD.

A regulation function is recommended to be collocated with the segregation point of an AD. By placing a regulation function with the FA segregation, the FIFO characteristic of the packets within an FA can be kept until the regulation.

9.1.2 Relay node capability

The relay nodes may have incomplete transport functionality. For example, a legacy node may not have the FA based scheduling or the regulation function. The FA based scheduling function is required to guarantee both 1) the FIFO characteristic among the packets within an FA and 2) the isolation of an FA with a separated queue. A simple FIFO scheduler with a single queue, as well as a weighted fair queuing scheduler with separated queues, would guarantee the FIFO characteristic for any FA. However, a FIFO scheduler can accumulate the maximum burst of FAs sharing the queue. If a cycle is formed by relay nodes with such FIFO schedulers, one cannot guarantee a latency bound. As such, it is required that the FA based scheduling supports the FA isolation.

Based on the supporting functions, relay nodes are categorized as the following:

- CAT 0: A relay node without the FA based scheduling or the regulation function.
- CAT 1: A relay node with the FA based scheduling but without the regulation function.
- CAT 2: A relay node with the regulation function but without the FA based scheduling

- CAT 2-1: A relay node dedicated for the regulation function. This type of relay node does not have the switching capability.
- CAT 3: A relay node with both the FA based scheduling and the regulation function.

The regulation refers to a function of keeping packets in a buffer according to a predetermined rule even if packet transmission is possible. It resides in relay nodes, or is placed in separated physical devices. It may be available only in some relay nodes. It is recommended that a regulation function is collocated with the segregation point of an AD. The location information of the regulation functions is recommended to be gathered prior to an AD design. This information is gathered through a dedicated interface with an automated procedure or manually. The regulation function, however, is independent of flow aggregation/segregation functions. The regulation functions may also be placed additionally at the middle of an AD.

A regulation function is categorized based on its queue management scheme and the regulation target entity. The per flow regulation function has queues per flow and regulates based on a flow-level regulation rule. The IR has a single queue for a set of flows but is based on a flow-level regulation rule. An IR examines the packet at the head of the queue, checks the flow it belongs to, and determines when to transmit the packet. The per FA based regulator has a single queue for the FA, and the regulation target is the FA itself. The regulation rule is based on the FA's parameters such as the sum of flows' arrival rates, which belongs to the FA. An example of FA-based regulators is described in Clause 9.3 and Appendix I.

9.1.3 AD decision & alteration

The ADs have to be determined with considerations of many aspects. The size of an AD decides the number of the boundary ports of the AD, therefore the number of input-output ports pairs of the AD, and the number of FAs within the AD. Smaller the AD, fewer FAs, fewer queues necessary, thus simpler the network schedulers. On the other hand, smaller AD means more ADs in the path, larger the latency bound. The balanced point in between has to be determined in the network configuration phase. As the network state dynamically changes, AD merge and division should also be possible.

The nodes' capabilities are also taken into consideration. CAT 2 or CAT 3 nodes with regulation functions, as defined in Clause 9.1.2, can be AD aggregation or segregation points.

Aggregation domains may be merged or divided anytime. The ADs have to be determined with considerations of many aspects. The size of an AD decides the number of the boundary ports of the AD, therefore the number of input-output ports pairs of the AD, and the number of FAs within the AD. Simply put, smaller the AD, fewer FAs, fewer queues necessary, thus simpler the network schedulers. On the other hand, smaller AD means more ADs in the path, larger the latency bound. The balanced point in between has to be determined in the network configuration phase or during the runtime. As the network state dynamically changes, AD merge and division is recommended to be possible.

9.1.4 FA granularity decision

One of the requirements for the framework is that the flows with different paths should belong to different FAs. Flows with the same path may or may not belong to the same FA. More criteria, such as latency bound requirements or TSpec of flows, can be considered for aggregation decisions. For example, only video flows of the same path are aggregated into a single FA. Finer granularity means better performance but more complexity. Since the granularity of FA is better to be consistent across ADs, in terms of latency bound performance, FA granularity negotiation among ADs is recommended to be possible.

9.2 Call setup procedure

The call setup procedure includes the admission control and resource reservation. Figure 5 describes the overall procedure for call setup.

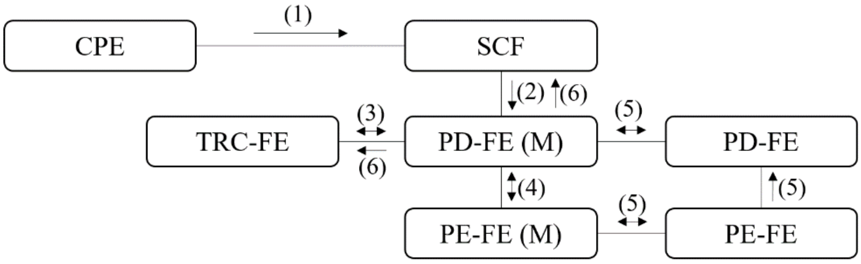

Figure 5 – Overall call setup procedure

1) CPE requests application-specific service to SCF with optionally a service QoS requirements.

2) SCF identifies flows corresponding to the CPE request, and extracts or derives the RSpec of the flows, then requests to PD-FE for the network configuration and admission control.

3) PD-FE requests and receives from TRC-FE for the network configuration information.

4) PD-FE requests and receives from PE-FE for the flows' enforcement information.

5) The master PD-FE requests and receives from the other PE-FE instances through PD-FEs on the flows' paths for the enforcement information. This exchange can optionally be done through PE-FEs.

6) PD-FE finalizes and notifies the admission decision result to TRC-FE and SCF.

There are two possible scenarios for passing the QoS information over an end-to-end path.

- In the first approach, the QoS requirements and information for a given flow's service can be passed over the end-to-end path through application layer signalling or through the Ri reference point between RACFs.
- In the second approach, the QoS requirements for a given service can be passed over the end-to-end path through path-coupled QoS signalling (e.g., resource reservation protocol (RSVP)-like [b-IETF RSVP]).

This Recommendation supports both approaches, as the RACF does. A customer premises equipment (CPE) may or may not have the signaling capability. A CPE may have signaling capability but may not be able to explicitly specify required QoS level. In such cases the SCF extracts or derives the TSpec of such CPEs and then requests admission of the flow and reservation of resources, from the RACF. The PD-FE is the final decision point of the admission.

The TSpec can include token bucket parameters (a burst size and an average input rate), a peak input rate (p), and a maximum packet size (M). If a packet is larger than M, then it is possible that the packet does not receive the same service with the conforming packets.

The SCF or the CPE can also derive the reserve specification (RSpec) of a flow, which can include latency upper bound information, and negotiate whether it can be met with the RACF. The QoS provisioning can be based on the network allowance. As a CPE specifies a flow's TSpec, based on the best end-to-end path among those that can be provided, the feasible latency bound is calculated and notified to the flow. The CPE decides whether to accept or not.

Existing best-effort service traffic should minimally affect the latency bound of the high priority flows. Networks should be aware of the best-effort service traffic and take it into consideration.

Interactions in the call setup procedure among supporting administrative domains shall be possible. The interactions between the networks at the transport stratum is through the reference point Ri in Figure 1.

In order to handle different types of CPEs and transport QoS capabilities, the RACF is required to support the following QoS resource control modes as part of its handling of a resource request from the SCF:

- Push mode: The RACF makes the admission and resource control decision based on policy rules and autonomously instructs the transport functions to enforce the policy decision.
- Pull mode: The RACF makes the admission decision based on policy rules and, upon the request of the transport functions, re-admit the resource request and responds with the final policy decision for enforcement.

9.2.1 Dynamic QoS negotiation

A static QoS negotiation, such as the IntServ, guarantees a fixed service rate to a flow during its lifetime. It is simple but may under-utilize the network. It is recommended that a dynamic QoS negotiation is possible. The SCF or the CPE can initiate the renegotiation. A single flow's renegotiation may result in all the other flows' renegotiation.

The service level negotiation can be a two-way handshake, or a more complex process. In a two-way handshake negotiation, the SCF or the CPE specifies its RSpec and TSpec. The RACF decides whether it can be met. If not, it denies admission.

Negotiations on RSpec and TSpec, such as the latency upper bound and the maximum burst size, are recommended to be possible, with more complex handshake processes.

The negotiation on both RSpec and TSpec at the same time can be effective. For example, the RACF can provide the CPE multiple {latency bounds, max burst size} pairs to choose. The CPE can select one of them. In general, a smaller burst size gives a smaller latency bound.

Exchange of dynamic admission control information, such as the current flow's latency guarantee status, is recommended to be possible. These exchanges of information can be periodic or on-demand.

9.3 Flow treatment procedure

After the call setup procedure (the admission and resource reservation process), an admitted flow shall be aggregated into a FA at the aggregation point of an AD according to the aggregation criteria applicable to the flow. Figure 6 describes the overall procedure for flow treatment.

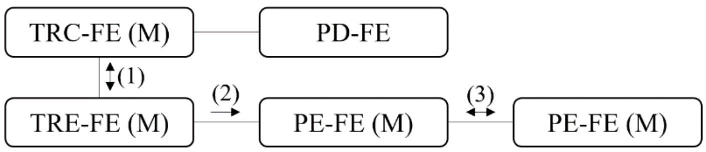

Figure 6 – Overall flow treatment procedure

1) After the call setup procedure is completed by PD-FE, TRC-FE(M) finally decides the ADs and FAs for the flows.

2) The master TRE-FE notifies the PE-FE of the updated FA enforcement policies and parameters.

3) The master PE-FE enforces the policies and parameters on FAs, and shares the information with other PE-FE instances on the flows' path.

The enforcement parameters, such as the allocated service rate and maximum allowable burst size, for each FA is determined at the TRE-FE and informed to PE-FE through reference point Rx. The queuing, scheduling, and regulation policy according to the FA parameter is determined at the PD-FE and informed to the PE-FE through reference point Rw. The PE-FE is the final enforcement point of the flow treatments.

An FA should be queued, scheduled, and regulated according to the requirements defined in Y.3113. In Y.3113 it is required that the FIFO characteristic of the packets in an FA should be preserved within an AD. A simple FIFO scheduler that accommodates all the FAs in a single queue, or separate queues per FA can fulfil the requirement. In order to minimize the E2E latency, it is recommended to have separate queues per FA and a fair scheduler for the queues at the output ports. A FIFO queue that accommodates all the FA in the same priority can be allowed, for relatively simple networks, in which a burst accumulation is not problematic.

Y.3113 also requires IRs per FA be placed at the boundary of an AD.

An FA within an AD is treated as a single control entity, i.e., the flow inside an FA is not a control target. However, whenever a flow joins/leaves, the schedulers and regulators should take these changes into account. The schedulers shall update the fair rates that should be allocated to the FA. The regulators shall update the proper sustainable rate and maximum burst of the FA to be regulated.

Based on the framework defined in Recommendation Y.3113, a network can have an arbitrary number of ADs with arbitrary sizes; and regulators in between.

An AD can be a single hop, i.e. is restricted within a relay node or spans over a single link. In this case a flow aggregate within the AD can then be configured to be a set of flows sharing the same input and output ports of the node. As such, some of existing QoS frameworks become a specific architectural example of the general framework described in this Recommendation. For example, IEEE time sensitive network (TSN) task group’s asynchronous traffic shaping (ATS) can be modelled with a strict priority scheduling node as a single hop AD and the minimal IR as a regulator. Another example is to model a strict priority scheduling node as a single hop AD and a regulator per FA, which is based on input/output port pairs of a flow. Such a regulator is called Port-based Flow Aggregate Regulator (PFAR). The details of PFAR architecture is described in Appendix I.

An AD can also be of multiple hops. The flow aggregates in such an AD is defined according to the ingress and egress ports of the AD. One possible configuration for an FA is to put all the flows with the same {ingress, egress ports} pair, thus having the same path, into a single FA. The critical design choice in the ADs with multiple hop is whether to allocate a separate queue in each node for an FA.

If so, then the scheduler for the queues of the FAs with the same priority should provide fair sharing, preferably be one of the fair-queuing schedulers, such as the deficit round robin scheduler.

If not, i.e. multiple FAs are put into a single queue, then the burst accumulations among the FAs occur and possibly the burst explosion happens as well because of the cyclic dependency. As such, careful planning to avoid the cyclic dependency is necessary. One way of avoiding such a problem is to place regulators inside the AD, to cut the cycles formed by FAs inside the AD.

10 Architectural consideration for IMT-2020 networks

In the IMT-2020 network and beyond there are inevitably multiple network domains, with possibly different QoS frameworks. For example, in the IMT-2020 networks, access networks, core networks, and the network slices ranging across core networks have different QoS provisioning architecture. The architectural aspect of the IMT-2020 network is considered.

The functional architecture depicted in Figure 1 can be mapped to the IMT-2020 network. An AD can be an access network (AN) or a core network (CN), or a part of a core network. Equivalently, a CN can be divided into several ADs. A RACF should control a single core network.

Since the basic control target entity in IMT-2020 is a protocol data unit (PDU) session, which is defined to be a set of IP flows, within the IMT-2020 networks the IP flows are aggregated to a certain degree. Aggregation of the PDU sessions having the same path of an AD is recommended.

In the IMT-2020 network framework [b-ITU-T Y.3102], the combination of session management function (SMF) and policy control function (PCF) has functionalities that cover Y.2111 RACF. SMF sets up PDU sessions for user equipment to control the user plane for PDU sessions' connectivity (e.g., selection/re-selection of user plane network functions and user path, enforcement of policies including QoS policy and charging policy). SMF gets policy information related to session establishment from the policy control function (PCF). PCF enables end-to-end QoS enforcement with QoS parameters (e.g., maximum bit rate, guaranteed bit rate and priority level) at the appropriate granularity (e.g., per UE, per flow and per PDU session).

The transport functions TRE-FE and PE-FE, specified in 8.1, are also required to be added to the IMT-2020 network framework, to SMF and PCF.

11 Security Considerations

The QoS management of IMT-2020 network includes user equipment, access networks, and core networks that are subject to security and privacy measures. Sensitive information should be protected as a high priority in order to avoid leaking and unauthorized access. Security and privacy concerns should be aligned with the requirements specified in [b-ITU-T Y.2701] and [b-ITU-T Y.3101].

Appendix I

Architectural example: Port-based FA Regulator (PFAR)

(This appendix does not form an integral part of this Recommendation.)

A port-based flow aggregate (PFA) is defined to be a set of flows with the same priority sharing the input and output port in a relay node such as a switch or a router. If there are N ports in a switch, and C classes, then there can be at most N^2C PFAs in the switch. There can be NC such PFAs in a single output port module, ignoring the fact that there is no flow having an output port that is the same with the input port. In this architecture a regulator may be placed for each high priority PFA in an output port module, just before the class-based queueing/scheduling system of the output port module. We call this regulator the Port-based flow aggregate regulator, PFAR. A PFAR sees a PFA a single flow with the parameter {the sum of initial arrival rates; the sum of initial maximum bursts} of the flows in the PFA, and regulates the PFA to meet the parameters. By the initial parameter of a flow, we mean the parameter of a flow at the source as it generates the flow according to a traffic specification (i.e. TSpec defined in DiffServ framework). The PFARs can be placed at the output port of a switch for the regulation of high priority traffic. Figure I.1 depicts an example architecture of the data-plane of a switch having the PFARs within the output modules.

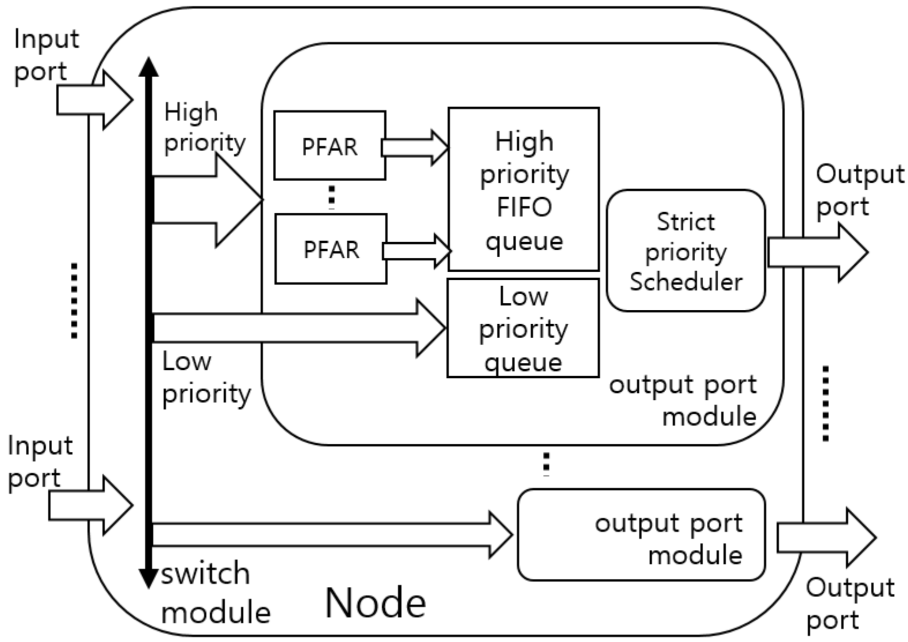

Figure I.1 – Example architecture of a node with PFARs at the output port modules

A network with switches with PFAR is an extreme example of the general architecture in Figure 3. Here the AD is from the scheduler to the next node's output module and the regulator is a set of PFARs in an output module of a switch, as depicted in Figure I.2. The AD encompasses submodules in two nodes, but the FA can be defined based on the second node's input and output ports.

Similarly, as another example, IEEE TSN ATS can be modelled with a strict priority scheduling node as a single hop AD and the minimal IR as a regulator.

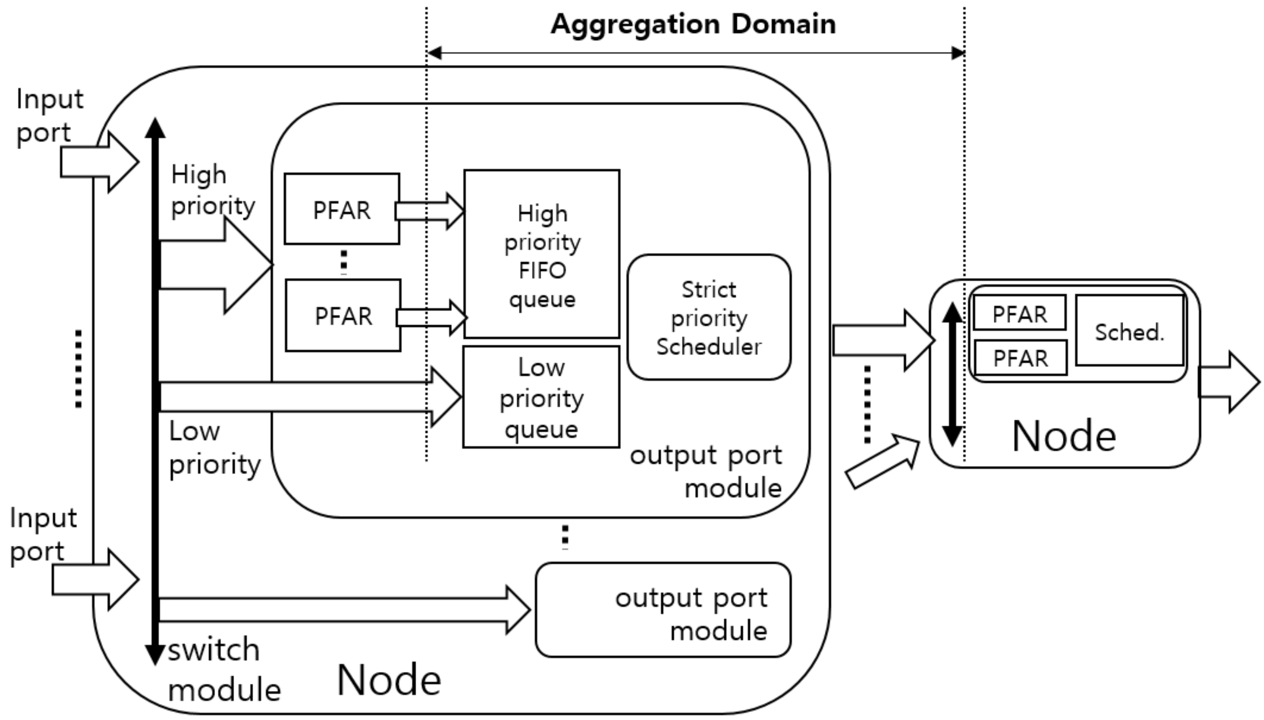

Figure I.2 – Mapping between the general architecture and PFARs at the output port

By PFAR, the complexity of regulation is reduced therefore the architecture becomes scalable. In ATS, two factors contribute to the implementation difficulty. First, it has to identify the flow that the packet at the head of the queue (HOQ) belongs to. The current flow state, thus the eligible time of the flow can then be obtained. Second, the individual flow state has to be maintained in order to be able to decide the eligible time of a packet. While the latter is the complexity within a control plane, the former impacts the real-time data-plane packet processing. With the PFAR, the HOQ flow identification process is unnecessary, and only hundreds of PFAs' states, instead of millions of flows' states, are maintained at a switch.

It is well known that a network with cycles suffers from the cyclic dependency problem. A carefully deployed PFAR, as well as an IR or a per-flow regulator, can break any cycle in a network. The delay bound of a network with PFAR is comparable to that of a network with ATS IRs. For detailed knowledge on the performance of PFAR, see [b-Joung-2022].

Bibliography

[b-ITU-T Y.2701] Recommendation ITU-T Y.2701 (2007), *Security requirements for NGN release 1*.

[b-ITU-T Y.3101] Recommendation ITU-T Y.3101 (2018), *Requirements of the IMT-2020 network*.

[b-ITU-T Y.3102] Recommendation ITU-T Y.3102 (2018), *Framework of the IMT-2020 network*.

[b-IETF RSVP] IETF RFC 2205 (1997*), Resource reservation protocol (RSVP).*

[b-Joung-2022] Joung, J., Kwon, J., Ryoo, J., and Cheung, T. (2022) *Asynchronous Deterministic Network based on the Diffserv Architecture*, IEEE Access, Vol. 10, Jan., 2022.
