

---

# White Paper

## MUSE QoS Architecture

---

Identifier: White Paper QoS  
Class: Report  
Version: 1  
Version Date: 10/01/2008  
Distribution: Public

## EXECUTIVE SUMMARY

This White Paper describes a comprehensive view of how Quality of Service (QoS) is handled within the MUSE architecture.

MUSE advocates the introduction of QoS into IP networks as this allows better resource utilization while at the same time it allows to serve multiple and different applications with the transport quality they actually need.

After analysing the failure of previous approaches to QoS, pragmatism, simplicity and cost-effectiveness are identified as the essential characteristics of a successful solution.

Apart from this, the solution needs to be able to offer quantitative QoS support for some services and qualitative for others, to support a retail/wholesale split in the QoS business model, to provide upstream QoS, especially across the access link, and to support multiple service edges.

Traffic classes, selective CAC and appropriate network dimensioning are the keystones of the solution proposed by MUSE.

The usage of at least four traffic classes (real-time, streaming, transactional and best-effort) is recommended by MUSE as the way for differentiating traffic whereas keeping the scalability of the network.

MUSE recommends a user-centric approach where classification of traffic into traffic classes is a responsibility of the user, although it is still expected to be often delegated to the providers. Traffic policing is recommended to be used for ensuring that the usage of each traffic class does not exceed what has been planned, agreed or contracted.

Classification of upstream traffic into traffic classes and its prioritization into the access link will be realized by the RGW according to user preferences. Prioritization of downstream traffic per traffic class is made by the network according to the rules described in user contracts.

MUSE recommends the use of central CAC for the small subset of services that actually need it (e.g. VoD) and only in those parts of the network where the network operator has identified a potential dimensioning problem. Local CAC at the Access Nodes or no CAC can be used for the rest of the traffic. Appropriate network dimensioning will help to minimize the risk of having congestion or blocking problems.

MUSE recommends a “provisioning” scenario where the central CAC, which has a view of all network resource usage, is able to allocate to a local CAC a certain amount of resources that will then be managed locally. Within this approach, the central CAC regularly monitors the usage of local resources at the Access Nodes and adjusts the resources allocated to the local CAC entities if necessary.

All the above recommendations coupled with a good network dimensioning form the recipe of the MUSE QoS solution.

## TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY</b> .....	<b>2</b>
<b>TABLE OF CONTENTS</b> .....	<b>3</b>
<b>ABBREVIATIONS</b> .....	<b>4</b>
<b>LIST OF CONTRIBUTORS</b> .....	<b>5</b>
<b>1 INTRODUCTION</b> .....	<b>6</b>
1.1 What is QoS about? .....	6
1.2 The need for a pragmatic solution .....	7
<b>2 ARCHITECTURAL DESIGN PRINCIPLES</b> .....	<b>8</b>
2.1 Architectural options .....	9
2.2 CAC options.....	10
2.2.1 <i>Central CAC</i> .....	10
2.2.2 <i>Local CAC</i> .....	10
<b>3 MUSE QoS ARCHITECTURE</b> .....	<b>11</b>
3.1 Architecture overview .....	11
3.2 Traffic classes.....	12
3.3 Selective CAC.....	13
3.4 Provisioning scenario for co-ordinating central and local CAC .....	14
3.4.1 <i>Central CAC implementation</i> .....	15
3.4.2 <i>Local CAC implementation</i> .....	15
3.5 Policy enforcement .....	16
3.6 Coordination between Home Network – Access Network for QoS .....	17
<b>4 ACCESS NODE AND RESIDENTIAL GATEWAY REQUIREMENTS FOR MUSE QoS</b> .....	<b>18</b>
4.1 Description of Access Node QoS functionality in the access .....	18
4.1.1 <i>Frame Treatment in the Residential Gateway</i> .....	18
4.1.2 <i>Frame Treatment in the Access Node</i> .....	20
4.2 Identification of requirements .....	20
4.2.1 <i>Functional Requirement on Residential Gateway</i> .....	20
4.2.2 <i>Functional Requirements on Access Node</i> .....	21
<b>5 CONCLUSIONS</b> .....	<b>23</b>

## ABBREVIATIONS

3GPP	3rd Generation Partnership Project
AN	Access Node
ASP	Application Service Provider
BNG	Broadband Network Gateway
BRAS	Broadband Remote Access Server
CAC	Call/Connection Admission Control
COPS	Common Open Policy Service
CPE	Customer Premise Equipment
DHCP	Dynamic Host Configuration Protocol
DiffServ	Differentiated Services
DPI	Deep Packet Inspection
DSCP	Differentiated Services Code Point
DSL	Digital Subscriber Line
EN	Edge Node
IGMP	Internet Group Management Protocol
IP	Internet Protocol
IPoATM	IP over ATM
IPTV	IP Television
ITU	International Telecommunication Union
L3	Layer 3
L4	Layer 4
MPLS	Multi Protocol Label Switching
NAP	Network Access Provider
NSP	Network Service Provider
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PE	Provider Edge
PIR	Peak Information Rate
RGW	Residential Gateway
SIR	Sustainable Information Rate
SLA	Service Level Agreement
SP	Strict Priority
TC	Traffic Class
TDM	Time Division Multiplexing
TE	Traffic Engineering
QoS	Quality of Service
VoD	Video on Demand
VoIP	Voice over IP
VPN	Virtual Private Network
WFQ	Weighted Fair Queuing

## LIST OF CONTRIBUTORS

Antonio J. Elizondo (editor)  
Friedrich Armbruster  
António Gamelas  
Govinda Rajan  
Pieter Nooren

Xavier Pougard  
Dave Thorne  
Henrik Villför

Telefónica I+D  
Nokia Siemens Networks  
Portugal Telecom Inovação  
Alcatel-Lucent  
Netherlands Organisation for Applied Scientific  
Research (TNO)  
France Telecom R&D  
British Telecommunications  
TeliaSonera

# 1 INTRODUCTION

The current best-effort approach used in the Internet does not differentiate traffic; it takes no account of the kind of application that has generated it or where the source or the destination points are located. In the best-effort approach, the normal practice is to dimension the network so that the network delivers a reasonable quality of experience, even at times of peak usage. However, this can lead to the necessity of almost continuously increasing the network transport capacity, with little or no revenue increase. This is a particular concern when a Service Provider has to pay for more peering services (interconnect) bandwidth.

MUSE advocates the introduction of Quality of Service (QoS) into IP networks as this allows better resource utilization while at the same time providing different applications (e.g. VPN, VoIP, IPTV, VoD, etc) with the transport quality they actually need.

This White Paper describes a comprehensive view of how QoS is dealt within the MUSE architecture.

In this chapter, first the different meanings of QoS are introduced. Second, a list of high level requirements for QoS is compiled once some of the problems of previous and current solutions have been identified.

The second chapter presents the most important architectural design principles and options followed by MUSE.

The third chapter describes the MUSE QoS architecture, starting from a high level view and going into further details of the main components: traffic classes, CAC, provisioning, policy enforcement, and coordination with the home network.

The fourth chapter identifies the main functionalities and associated requirements that must be present at residential gateways and access nodes.

Finally, last chapter concludes with a summary of the main features of the QoS solution proposed by MUSE.

## 1.1 What is QoS about?

QoS is sometimes thought of as a collection of packet forwarding performance metrics such as packet loss, mean delay, jitter, throughput, etc. However, there are other interpretations that include service or network availability aspects. An initial common understanding of the meaning of QoS is needed before describing a QoS architecture.

QoS is defined in ITU-T E.800 as “the collective effect of service performance which determines the degree of satisfaction of a user of the service”.

From this point of view, QoS is defined from the end users perspective, and the problem of providing QoS consists of satisfying users. Users are satisfied when their service perceptions are aligned with their former expectations. The problem is that these perceptions are based on both subjective and objective views, with the subjective part being difficult for providers to control. Because of this, QoS approaches are usually constrained to the problem of providing the objective part of the problem, i.e. network performance.

Note also that QoS can be described in quantitative or qualitative terms. In the case of a quantitative approach, bounds on statistical or deterministic parameters are established and contracted in an SLA. The problem then becomes one of monitoring the SLAs to demonstrate whether or not these bounds have been met. The qualitative approach is generally simpler, and often associated with relative (as opposed to absolute) traffic treatment. It typically has no associated parameters, but must deliver sufficient user experience benefit to justify using QoS.

## 1.2 The need for a pragmatic solution

Providing QoS in IP networks is a problem which could be solved in a variety of ways. However there is still no consensus on the best way. Some solutions (e.g. IPoATM, IntServ, MPLS TE, ...) are rather complex to be cost-effectively deployed in current access and aggregation networks, as they require step changes both in network equipment and their control and management, whereas others (e.g. pure DiffServ only) may seem to be too simple or insecure to be attractive to network operators.

Today's DSL broadband access architectures (based on the DSL Forum TR-101 architecture) consist of aggregating connectivity for mass-market Broadband Access in which all traffic goes to a single aggregation point, the BRAS. In fact TR-101 slightly extends this model, to allow for a second BRAS or BNG, which is specifically included to allow video content to be sourced from a separate location. The QoS control is however centred on the primary BRAS. Bandwidth control is limited to the downstream direction, and is based either on the concept of hierarchical scheduling or per hop behaviour.

Part of the problem is whether or not to try and replicate the traditional carrier class quality associated with TDM-based leased lines. Another difficulty is the multiple types of different services that IP networks need to support. However, one of the biggest barriers to the introduction of QoS support is the difficulty of appropriate configuration and management, especially for the residential mass market (simply due to its large scale).

MUSE therefore advocates a pragmatic and simple way to provide services with some degree of QoS. Pragmatism is needed to design a solution scalable enough to address the residential market and flexible enough to address the corporate one. By focusing on the needs of triple play offerings, it is possible to design a simple and cost-effective way to provide sufficient QoS for the mass market. QoS will be a way to improve value of services that are offered by the network operator and thereby differentiate them from similar services offered over the best effort Internet. However, the proposed solution should have a wider applicability than just triple-play, in order to adapt to new residential services or corporate market needs, although it may not meet the requirements of any conceivable service.

Furthermore, triple play services demand high bandwidth that will require big investments from the Telecom Operator. These could be financed from revenues from the users, but there are signs that they are quite reluctant to pay any more money only for plain connectivity services (i.e. Internet). However, if the operators are able to differentiate their connectivity resources, they can better sell them to connectivity customers no matter if they are Services providers, Packagers, Enterprise users or Residential users with specific requirements.

The key issue is to provide a solution whereby operators are able to virtually segregate their networks in connectivity resources with specific QoS that can be offered on-demand to "connectivity resource" consumers. This business model can lead to different architecture scenarios where the user, the Service Provider, the Packager or all of them request resources with QoS for certain applications.

The requirements for QoS can be summarized as follows:

- The need for a pragmatic, simple and cost-effective approach for delivering mass market services with an acceptable QoS from the point of view of both the customer experience, and the willingness of the service provider to pay for the additional network functionality. The QoS for these services may simply be non parameterised traffic differentiation.
- The need to be able to offer quantitative QoS support for some services
- The need to support a retail/wholesale split in the QoS business model
- The need to provide upstream QoS, especially across the access link, to support services like quality VoIP, video telephony, video conferencing, etc
- The need for supporting multiple service edges. There are several different types of service edges in addition to a BRAS, for example video servers, soft switches, and PE routers. Typically the services associated with these Edge Nodes<sup>1</sup> do not go via the BRAS (for scalability and reliability reasons), and so there is no longer a single point of QoS control.

## 2 ARCHITECTURAL DESIGN PRINCIPLES

The MUSE QoS architecture is based on the following principles:

- As already argued, the first and probably most important is that the solution should be simple and pragmatic. Scalability and management are the main concerns when addressing the mass market.
- The solution should have sufficient flexibility to allow it to be adapted for the corporate market.
- The solution should be designed to be applicable to the wholesale model, so that a third party is then able to offer a similar service level in the retail model.
- Not all traffic requires hard, parameterised QoS guarantees. Only a minority of the traffic will require such hard guarantees. The rest of the traffic will have softer requirements, and hence may be handled with simpler mechanisms, or indeed no QoS requirements. Note that today's mass market services are generally not guaranteed with regard to either performance or availability. There are no Service Level Agreements (SLAs) between residential users and service providers, who just aim to ensure that user perceptions of the offered services match expectations.
- There is not necessarily a single solution that is valid for every network segment (access, aggregation, and core). The characteristics of the different types of segments are very different, so a solution that is optimized for one domain will probably be (very) suboptimal for other domains. In general, more complex QoS mechanisms are needed when resources are scarcer. For instance, when resources are very scarce (e.g. wireless media), dynamic resource reservation may be appropriate. However, QoS control per individual flow would be overkill in core networks, and would lead to difficult scaling problems.

Based on these principles, MUSE advocates a global approach to solve the QoS problem. It is not only to design QoS control mechanisms but to follow a set of good practices that should be followed on a continuous basis for providing QoS.

---

<sup>1</sup> Edge Node is the MUSE terminology for those IP nodes that are situated at the edge of the aggregation network with other networks. It is equivalent to the BNG defined by the DSL Forum.

The steps involved in deriving and delivering an appropriate QoS solution are as follows:

- Identifying what the **user required QoS** is, by analyzing current service expectations, but also taking into account market trends and emerging applications.
- Deciding what the **provider planned QoS** should be, taking into account business objectives and techno-economical constraints, and providing the necessary network mechanisms appropriately configured to make the job
- Assessing what is the actual QoS the **provider delivers** (by means of monitoring) and what the **user perceived QoS** (by means of obtaining feedback from customers) are, and taking the necessary actions to correct the deviations from the provider planned QoS.

Note that definitions of user required QoS, user perceived QoS, provider planned QoS and provider delivered QoS can be found in ITU-T G.1000.

## 2.1 Architectural options

There are some basic questions that need answering before defining the MUSE QoS solution:

### **Q1: How to differentiate traffic?**

MUSE recommends simple traffic differentiation (i.e. a la DiffServ) as one of the basic QoS techniques. However the differentiation should be done in such a way that it does not impose a heavy burden in terms of complexity and performance. Therefore, MUSE advocates using a limited set of traffic classes (e.g. four classes) inside the access and aggregation networks.

### **Q2: Who is responsible for the classification of the traffic?**

The nature and importance of the traffic should be taken into consideration in assigning the responsibility to classify traffic, and these aspects are better known to the users of the network (Residential, Corporate, Service Providers, Packagers, etc). Indeed, classification of traffic into traffic classes should be done before the ingress point of the network acknowledging the right of the users to decide this and this is the only way to guarantee QoS on all the network segments.

However, as most users are laymen and do not know how or not want to make this kind of decisions, it is expected that most of them would delegate these decisions to the network as part of the function of users' preferences. Note also that, whereas it is legitimate to try to offer this capability as an added value so that network users can be liberated from taking decisions, it is also true that some users could prefer to keep the control, and hence this model should not be imposed but offered as a possibility.

### **Q3: How is QoS requested?**

Requesting QoS, once a DiffServ approach has been selected, is implicit to the fact of belonging to a given traffic class, as the network will normally define a set of constraints (e.g. a given maximum mean bitrate per traffic class) for using the different available traffic classes on a per user basis. This will basically depend on the physical constraints of the access technology and on the contract negotiated with the user. Verifying that the constraints for using the traffic classes are being fulfilled by the users is part of the policy enforcement, and this should ideally be done at the ingress points of the network. Note that these constraints are expected to be rather static in time, although they may be altered upon previous negotiation between the user and the network.

Additionally, explicit guarantees per session can be provided to some services by means of a CAC (Call/Connection Admission Control) system so that service session requests are accepted or rejected based on the availability of requested network resources for the traffic class applicable to the session. Such a CAC system will receive explicit requests through application's own signalling (e.g. in the Session Description Protocol) or from application servers.

#### **Q4: How are the resources provided?**

The goal of CAC is to verify in real time that there are available resources enough for satisfying the QoS guarantees. In this sense, admission control has a complementary role to network dimensioning, which should have provided the necessary resources in advance. Note that QoS should be largely achieved by the correct network dimensioning. Even with CAC, if the network is largely underprovisioned, a significant number of session requests will be denied, thereby leading to a poor customer experience. CAC is therefore not a substitute for good network dimensioning, which is the basic means of providing the necessary resources for QoS support.

## **2.2 CAC options**

### **2.2.1 Central CAC**

A central CAC system is where all CAC decisions are made at the same place. The central CAC system has a complete view of the resources of the appropriate parts of the network. All call admission requests have to be signalled to this central system. For each and every call (signalled) request or (non-signalled) detection, the central CAC system is consulted, which, on the basis of resource availability, decides whether to allow or block the requested call. This decision is sent to the boundary node where enforcement may be done.

The main advantage of a central CAC system is that it is simple to manage. However, it requires the exchange of signalling messages, which may compromise the scalability and adds delay in the decision process.

### **2.2.2 Local CAC**

A local CAC system is where uncoordinated CAC decisions are made at the appropriate nodes on the basis of the state of a local link. For a CAC decision to be made locally and independently of other nodes, it is necessary that there be a local view of the availability of network resources that can be used by that node. To do that, it is necessary to partition network resources for each traffic class and to allocate them to the different nodes.

Resource partitioning naturally leads to a reduction in the potential multiplexing gains as unused resources allocated to a given node can not be used by another node or by another traffic class at the same node. In order to avoid this, partitions should be updated on a regular basis by a central authority that has a historic and global view of network resource usage or when a threshold associated to a given allocation is exceeded. Any re-allocation of resources should take into account the commercial agreements (SLAs) between network operators and service providers, which might specify bandwidth on a link-by-link basis.

A local CAC system reduces both the time to make decisions and the exchange of signalling messages. It can take advantage of direct interaction with IGMP messages, hence offering a way of implementing CAC for multicast traffic. However, its implementation is complex and must be carefully done in order to avoid problems such as the lack of information consistency.

### 3 MUSE QOS ARCHITECTURE

#### 3.1 Architecture overview

A diagram of the MUSE QoS architecture is described in Figure 1.

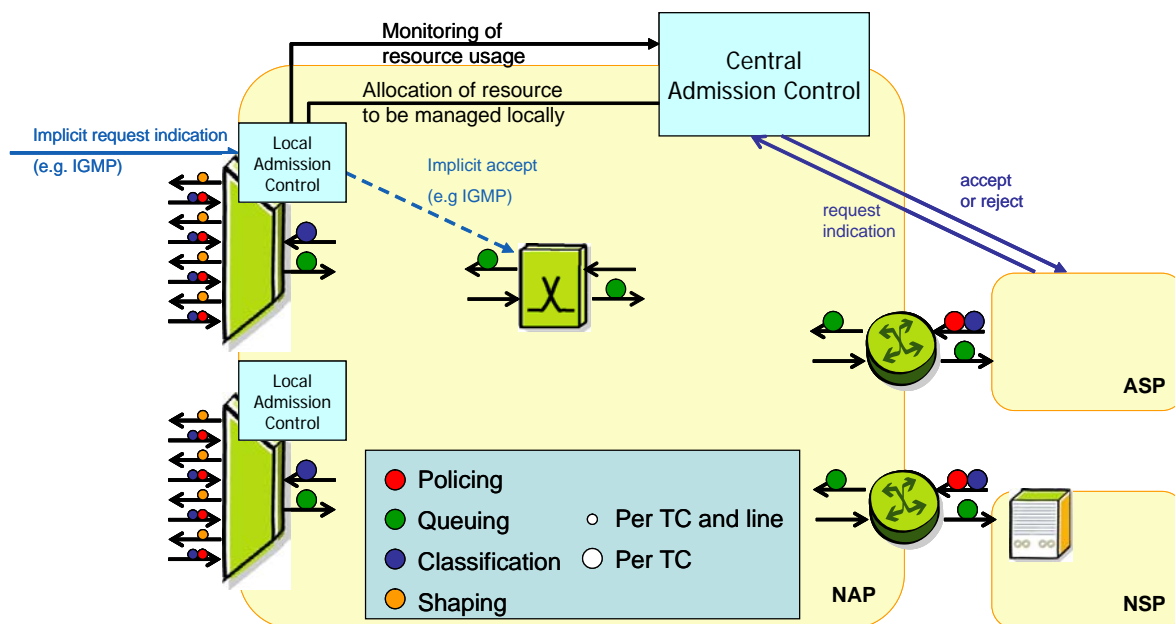


Figure 1: MUSE QoS Architecture

Traffic classes, selective CAC and appropriate network dimensioning are the keystones of the solution.

While traffic class differentiation will be used for most of the traffic keeping the solution simple and scalable, per flow differentiation can also be supported for certain types of traffic in the appropriate parts of the network. A small set of traffic classes will be used to deal with most of the traffic, establishing simple scheduling algorithms in the network elements to differentiate them.

Traffic classification will normally be performed by the traffic originator, unless delegated to the network operator. However, the network will ensure that the load of each traffic class is below the level that has been used for dimensioning by means of policing the different traffic classes on a per subscriber basis. This traffic policing should be done as close to the user as possible in order to react faster, to guarantee enough network resources, to protect sensitive traffic, etc; that is, at the access nodes. The policies could be statically provisioned for common residential services. However, personalised policies could be applied in association with the authentication process, so that richer and more dynamic services could be supported. Inside the network, nodes must have dedicated queues per traffic class at the output ports, which will be attended by a scheduling mechanism with strict priority or weighted mechanisms.

CAC is not necessary for all services, but only for those featuring difficulty in demand prediction and require a large amount of resource. This is likely to happen on the access links, because of their relatively high associated cost, and especially for the upstream traffic when using asymmetric access technologies. It may also happen on aggregate links which carry a mix of basic Internet traffic and video. Hence, CAC is recommended to be applied at least for protecting the traffic of the most sensitive services; that is, real-time or loss-sensitive ones, and especially those that require a relatively big amount of bandwidth (e.g. IPTV and VoD).

Appropriate network dimensioning will reduce the chances of having congestion problems and hence will diminish the need for widespread deployment and usage of resource reservation and admission control mechanisms; these can be difficult to deploy as they add the need for signalling and stateful management at network elements. Of course, an appropriate network dimensioning requires an awareness of possible traffic evolution patterns in the network. Hence, traffic monitoring and reporting per node are required.

## 3.2 Traffic classes

A natural way to group telecommunication services is as a function of the type of traffic they generate. The main differentiators identified in MUSE are:

- **Elasticity level (elastic/inelastic):** Elasticity level refers to the level up to which the traffic's original shape can be modified. Not all applications have the same elasticity level. Normally, communication services aspire to keep both data and temporal integrity. In order to establish the elasticity level of a given service/application, it is useful to assess which of both integrities is more important. So, elastic and inelastic applications (or traffic generated by those applications) can be distinguished as a function of which of these is more relevant. If data integrity is more relevant (e.g. a file transfer), lost or corrupted data have to be retransmitted, the traffic is considered elastic. If temporal integrity is the main concern (e.g. a voice call), there is normally no chance of retransmitting lost or corrupted data, so this type of service is characterised as having inelastic traffic. For some services both data and temporal integrity are important (e.g. certain interactive games).
- **Interactivity level (interactive/non-interactive):** The interactivity level describes the time integrity in both directions of communication. For instance, elastic traffic with a high interactivity level is generated by an application where both data integrity and temporal integrity are very relevant (for instance, e-commerce, etc)

- **Service availability (standard/high):** Availability is a very important consideration, and of course must be used as an attribute to identify the different traffic classes. Indeed, it is one of the most important considerations to take into account in core networks (which are not in the scope of MUSE).

The following table provides a mapping between these concepts and the ITU and 3GPP terminology.

Traffic class		Terminology proposed in MUSE	3GPP	ITU
Elastic	Non-Interactive	<i>Best effort</i>	Background	Non-critical
	Interactive	<i>Transactional</i>	Interactive	Responsive
Inelastic	Non-Interactive	<i>Streaming</i>	Streaming	Timely
	Interactive	<i>Real Time</i>	Conversational	Interactive

Table 1: MUSE traffic classes

It is recommended that all network nodes support at least the 4 classes of services defined in Table 1. Every operator can obviously provide additional classes of services.

Inside the network, nodes will have dedicated queues per traffic class at the output ports, which will be served by Strict Priority (SP) and/or Weighted Fair Queuing (WFQ) scheduling mechanisms.

### 3.3 Selective CAC

In large access network domains, there could be a scalability problem when implementing signalled and central CAC for each and every flow. Hence, MUSE recommends the use of central CAC for the (small) subset of services that actually need it, and only in those parts of the network where the network operator has identified a potential dimensioning problem, whereas local CAC or no CAC can be used for the rest of the traffic.

A mechanism is needed to segregate the network resources into:

- A set of resources that can be used by services that need no CAC, and are policed only on a traffic class basis so that a maximum class bandwidth cannot be exceeded.
- A set of resources for services that are controlled by a central system on a per call/session basis with explicit signalling.
- A set of resources for services that are controlled by a local CAC system on a per call/session basis with or without explicit signalling.

The next section describes the mechanism selected by MUSE to segregate the network resources.

Note that a given set of resources can either be completely dedicated to traffic subject to CAC or, to improve network utilisation, may also be shared by traffic not subject to CAC. In the case of sharing, prioritisation mechanisms will be required in addition to CAC. Depending on the network architecture and dimensioning, CAC may only be needed for certain links within an end-to-end path.

In addition to any admission control performed by the network operator, it is the responsibility of the service provider, through a separate service admission control, to check whether necessary resources are still available on the services platform and at the traffic classes contracted to the network operator (in a wholesale scenario).

### 3.4 Provisioning scenario for co-ordinating central and local CAC

MUSE recommends a “provisioning” scenario where the central CAC, which has a view of all network resource usage, is able to allocate to a local CAC a certain amount of resources that will then be managed locally.

Simpler scenarios could be designed where each CAC system just controls a dedicated set of resources and no interrelation is needed between them. However, the main drawback of such an approach is that optimization of resource usage is difficult to achieve.

Within this approach, the central CAC regularly monitors the usage of local resources at the Access Nodes and adjusts the resources allocated to the local CAC entities if needed.

The usage of WFQ is recommended as scheduling algorithm in those links where capacity is high enough, as it helps to provision the required bandwidth for each traffic class by appropriately setting the different weights. However, use of SP is recommended for guaranteeing a low queuing delay for real time traffic in those links where capacity is relatively low, whereas WFQ can be used for sharing the remaining bandwidth among the rest of classes.

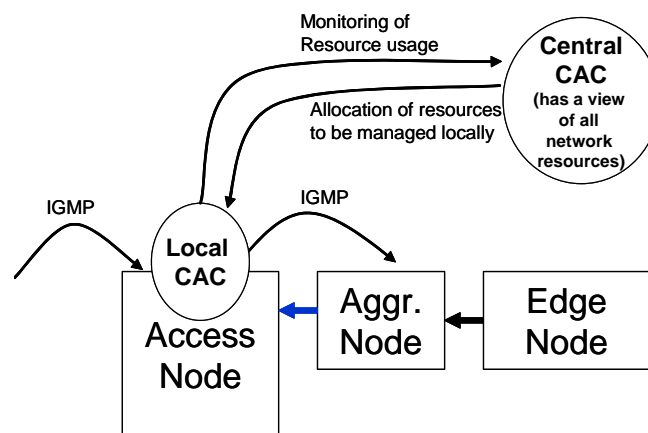


Figure 2: Provisioning scenario

This approach is recommended as it gives more flexibility to share resources between central and local CAC, and allows more reactivity if a significant evolution occurs between central and local traffic proportions. This mechanism could be used to adjust the threshold for local resources over a day, allowing, for instance, attractive prices on VoD when TV bandwidth is not heavily used, or even close to real-time when local allocated resources are exhausted while there are still global available resources, so that resource allocation is kept to an optimum.

More complex and dynamic scenarios can be envisioned where local CAC entities are able to request additional resources to the central CAC system in a proactive way. However, the added complexity does not justify going for such approaches in the medium term.

### 3.4.1 Central CAC implementation

Central CAC is usually considered as the simplest approach, and hence is in principle recommended to be used for applications requiring explicit CAC (e.g. VoD) where neither scalability nor setup delays are special concerns.

A central CAC entity has a single but complete view of the availability of all network resources in a network area. This is usually done by listening to link-state routing protocols running on the network, in conjunction with a database containing the installed network elements and the installed bandwidth per interface. The admission control entity receives all the requests for starting a service and, since it has a current view of the status of the network, it decides whether it is possible or not to accept the request. Once this decision has been made, resource reservation can be done.

However, for the admission control of multicast flows, a centralized implementation of the CAC function is not appropriate. A centralized admission control would have to decide if a new requested channel could be delivered or not each time it receives a request to join a multicast channel. There are three problems with this:

- The first is the sheer volume of requests which arise when people are channel zapping.
- The second problem is that the multicast protocol (i.e. IGMP) has no mechanism (i.e. no parameter in the message) to convey the required bandwidth, which means that there has to be a local association between bandwidth and channel.
- Finally in most cases the stream join will occur (automatically) closer to the end-user than the location of the admission control system, so there is no mechanism for the CAC system to actually prevent the join.

Two of these problems can be solved by a distributed CAC system, but there will still be a need to maintain a mapping of multicast group addresses to channel bandwidth.

### 3.4.2 Local CAC implementation

Local CAC allows to significantly reduce complexity, signalling exchanges and time required to provide a CAC answer.

By using traffic classes, it is possible to decouple the checking of available resources into two parts: a CAC per traffic class, which consists of verifying that the traffic class involved in the session request has available resources, and a policing of traffic classes.

Note that local CAC can be performed independently in different nodes (i.e. RGW, AN, EN), so that a single reject decision is enough for denying the service request. In this way, local CAC can be deployed at those parts of the network where congestion is more likely to appear.

Local CAC can easily handle multicast traffic (e.g. IPTV), as it is possible to be locally aware of IGMP messages.

This decoupling approach allows a positive local CAC decision made by a single node to be enough for guaranteeing that there will be available resources in the corresponding traffic class, provided that such a traffic class is dimensioned to be used at the maximum rate by every user. This is not a special concern for multicast traffic, as this does not depend on the number of users but on the number of channels that are being simultaneously watched. But this approach could also be followed for unicast traffic. By taking into account the statistical nature of the traffic generated by the services of a given traffic class, the amount of resources needed is reduced and scalability is improved.

In addition, under congestion conditions, the network operator may decide to unilaterally vary the policing of the traffic classes, communicating to the local CAC entities the new limits for each traffic class so that the new conditions for using the traffic classes can be taken into account. Besides, traffic policing could be dynamically varied upon request from the network users provided that there are enough available resources in the network.

Note that local CAC can be applied at the Residential Gateways, independently of other CAC systems running in the network. This way, services that are not subject to CAC in the network could be delivered with QoS guarantees without the network being aware of such service sessions. However, the Residential Gateway should have awareness of those service sessions, by snooping the application signalling or by using service signatures provided by the service providers. This last approach is better, as it requires the cooperation of the service provider, which can provide also the parameters needed for evaluating the CAC, i.e. the effective bandwidth. It can also provide the necessary signalling messages for informing the service provider about the denial of service and/or the user when the Residential Gateway CAC decides to block the session.

### 3.5 Policy enforcement

A policy is the combination of rules and services where rules define the criteria for resource access and usage.

According to the Common Open Policy Service (COPS) terminology, three functional elements are defined for using policies in a network:

- 'Policy Repository', which contains the policies that have to be applied in the network.
- 'Policy Decision Point' (PDP), which evaluate the policies upon a given request and notifies the decision to the corresponding Policy Enforcement Points. A CAC decision is an example of policy decision.
- 'Policy Enforcement Point' (PEP). The PEP is the place in the network where the policy decisions are actually enforced (e.g. access control and traffic policing).

Traffic policing consists of verifying that a given traffic class does not exceed a certain profile. Enforcement of the allowed QoS policy (bandwidth, maximum size of packet, etc.) maybe required so that misbehaving traffic does not impact the QoS of the other users/classes. Protection can only be provided if such enforcement is done on a per user and per traffic class basis. Note that as a result of the policy enforcement, out of profile traffic could be dropped, remarked or delayed until it complies to the profile (i.e. traffic shaping).

For the upstream traffic, it is recommended that the policy enforcement be done at the Access Nodes. This will minimise the chances of misbehaving users altering the QoS of other users in the aggregation network. Otherwise, excessive traffic marked as high priority by some users may cause starvation of lower priority traffic of other users.

However, for the downstream traffic, it is recommended that policy enforcement be done at the Edge Nodes on aggregate of IP flows (i.e. traffic classes) in order to lower the processing power required at these nodes. This will not prevent a single misbehaving user from impacting the aggregate, but will limit any damage to that aggregate.

In addition to the aggregate enforcement at the Edge Nodes for downstream traffic, it is recommended to have per user shaping of downstream traffic at the Access Node to prevent congestion in the first mile. Shaping is commonly done at the BRAS in current architectures. However this is no longer viable in a multi-edge architecture. The only point at which all the traffic for a given line comes together may be the Access node itself.

### **3.6 Coordination between Home Network – Access Network for QoS**

The link between the Residential Gateway and the Access Node is usually the main bottleneck, especially since its capacity cannot easily be increased. To solve the contest for bandwidth on this link, the usage of prioritization mechanisms for handling traffic classes is recommended. Additionally, traffic policing can be used for avoiding starvation of lower priority traffic and fulfilling the usage limitations per traffic class.

Classification of upstream traffic into the traffic classes, and its prioritization onto the access link will be realized by the RGW according to user preferences (note that these preferences can be delegated to the network operator or to any service provider, and that traffic could have been previously marked by a terminal). This classification can be done by identifying application signatures, by using predefined ports at the RGW, by evaluating the Ethertype, etc. The policing of upstream traffic will be done by the Access Node according to the rules defined above.

Prioritization of downstream traffic per traffic class is made by the network according to the rules described in the user contracts (i.e. residential and corporate users, service providers, etc). This is currently done at Edge Nodes. However, when having multiple Edge Nodes in the network, there is not a single point of control of downstream traffic except for the Access Node.

## 4 ACCESS NODE AND RESIDENTIAL GATEWAY REQUIREMENTS FOR MUSE QoS

### 4.1 Description of Access Node QoS functionality in the access

The Figure 3 shows a diagram of the different functionalities associated with QoS used at the Access Node and the Residential Gateway<sup>2</sup>.

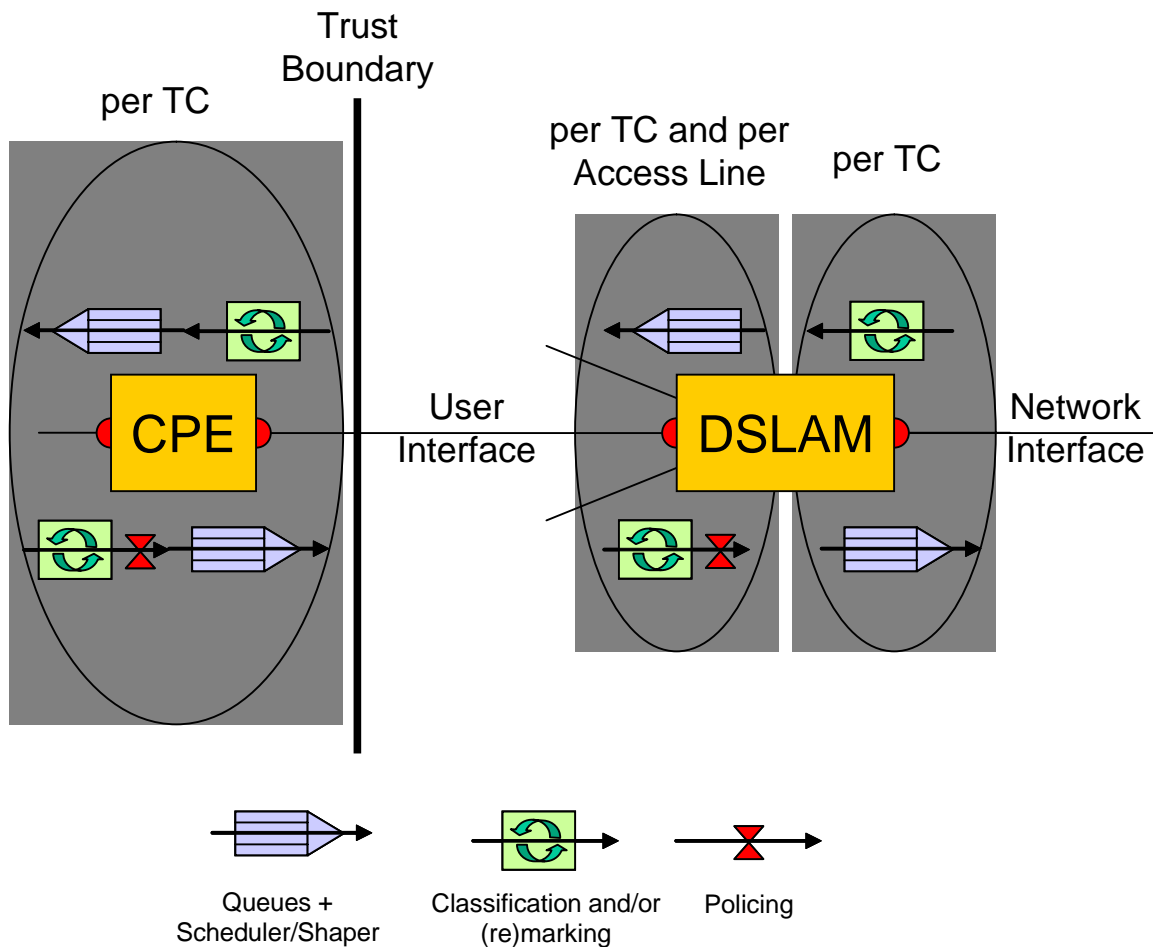


Figure 3: Access Node QoS functionality

#### 4.1.1 Frame Treatment in the Residential Gateway

##### Upstream Traffic Description

The functionality is described by following a frame travelling from the customer terminal through the Residential Gateway to the access link. One should keep in mind that the RGW will usually have a role of a switch connecting different terminals within the home network. Local flows are not considered explicitly in the description below but one should expect requirements on support of traffic differentiation also for flows within the home network.

<sup>2</sup> CPE is just shown for completion purposes. However, the description of QoS functionalities associated to the CPE/RGW is not exhaustive.

1. First the frame needs to be classified into a recognized Traffic Class:
  - No operation is done if the frame is marked with a recognized traffic class or traffic classification has not been delegated to the Residential Gateway.
  - A remarking and/or encapsulation operation takes place if the subscriber classification does not fit into the traffic class scheme used on the access link. This could be the case if the packet is marked with a traffic class that is not used for upstream traffic on the access link or if the marking is done at another protocol layer on the access link compared to the home network.
  - There may be a complex operation that includes analysing the type of traffic (e.g. by analysing L3/L4 headers) and adding the correct traffic class mark since the Residential Gateway is responsible for the correct marking of traffic that enters the access network.
2. Optionally, traffic can be policed in order to limit the traffic per Traffic Class to the allowed amount. Traffic policing may result in remarking or dropping of frames in case of excess traffic.
3. Then the traffic is forwarded to the access link interface and handed over to the correct outgoing queue.
4. Finally the outgoing scheduler selects frames from the queues according the defined scheme (e.g. SP or WFQ) and sends them out on the access link.

### **Downstream Traffic Description**

In the downstream direction (from the access link onto the home network) policing is not expected to be performed since there are no contractual motivation for it and since the access link (that has already been passed) is expected to be the bottle neck in most cases:

1. First the incoming frame is to be classified:
  - No operation is done if the frame is marked in the appropriate way for the home network.
  - Remarking and/or de/encapsulation operation if the classification does not fit into the traffic class scheme used on the home network. This could be the case if the packet is marked with a traffic class that is not used on the home network or if the marking is done at another protocol layer on the access link compared to the home network.
  - There may be a complex operation that includes analysing the type of traffic (by using DPI techniques or just by analysing L3/L4 headers) and adding the correct traffic class mark.
2. Then the traffic is forwarded to the home network interface and handed over to the correct outgoing queue.
3. Finally the outgoing scheduler selects frames from the queues according the defined scheme (e.g. WFQ or SP) and sends them out through the access link.

### **Local CAC in the Residential Gateway**

It may be of interest to implement some form of local CAC in the Residential Gateway to optimise the use of the access link. This could introduce additional steps in the flow description above such as snooping of signalling traffic or load monitoring. Note that this is not expected to imply dynamic interaction with the network provider and thus no signalling interface is required.

#### **4.1.2 Frame Treatment in the Access Node**

##### **Upstream Traffic Description**

The functionality can be explained by following a frame travelling from the User Interface (Trust Boundary) to the Network Interface:

1. First the frame needs to be classified into a recognized Traffic Class:
  - Classification is not done if the frame is marked with a recognized traffic class and traffic classification has not been delegated to the network.
  - Remarking can occur if the subscriber classification does not fit into the network traffic class scheme (Mapping must be configured)
  - Classification may be a complex operation that involves DPI techniques or a simpler analysis of the L3/L4 headers.
2. The Second step is to police traffic in order to limit it to the allowed amount per Traffic Class and per Access Line. Traffic policing may involve just drop precedence marking or the actual dropping of frames.
3. Traffic is forwarded to the network Interface and handed over to the correct outgoing queue (where it is mixed with the traffic from other subscribers with the same traffic class).
4. Finally the outgoing scheduler selects frames from the queues according to the defined scheme (SP or WFQ) and sends them out on the NNI interface.

##### **Downstream Traffic Description**

A similar sequence happens in the downstream direction, with the difference that no policing per traffic class is required at the network ingress interfaces of Access Nodes, since each traffic class should be under control by a combination of downstream traffic policing at the Edge Nodes and proper network engineering. Also, and depending on the Service Providers requirements, marking the traffic as Best Effort may happen at the egress point.

## **4.2 Identification of requirements**

The following requirements specify, in more detail, what MUSE Access Nodes and Residential Gateways must support.

### **4.2.1 Functional Requirement on Residential Gateway**

#### **General requirements**

- The Residential Gateway must be able to mark the traffic according to the network policies.
- A Residential Gateway should support at least four traffic classes inside the home network as well as between the home and the provider network.

#### **Requirements at the home network interfaces**

- The Residential Gateway must support separate queues for the traffic classes serviced by a weighted scheduler or priority queuing for downstream traffic.
- The Residential Gateway must support (re)marking of traffic forwarded to the access link.

#### **Requirements at the access link interface**

- A Residential Gateway must support queuing per aggregated (upstream) traffic class at the port connecting the access link. The queues should be serviced with a weighted scheduler or strict priority.
- The Residential Gateway should support per traffic class policing of the traffic towards the provider network.

#### **Requirements for configuration**

- At the access link port, the QoS Profile that defines the QoS Values per aggregated Traffic Class for the upstream traffic must be configurable. These profiles may be dynamically adjusted.
- The Mapping Table from e.g. application signatures, 802.1p or DHCP Bits of upstream traffic to the Traffic Classes supported on the provider network must be configurable. This mapping may be dynamically adjusted.

#### **Requirements pertaining to Call Admission Control**

- If a Call Admission Control function (CAC) is implemented it must be designed to function as a local CAC with a configurable policy.

#### **Management interfaces to the Residential Gateway**

- There should be a management interface to the RGW for monitoring and provisioning.

### **4.2.2 Functional Requirements on Access Node**

#### **General Requirements**

- An Access Node must support at least 4 traffic classes.
- An Access Node must be able to use 802.1P bits or DSCP Bits to classify traffic.
- An Access Node must support the remarking of received 802.1P or DSCP bits to traffic classes supported by the network.

### Requirements at the User Interface

- An Access Node must support queuing and shaping per (downstream) traffic class at the subscriber egress ports.
- An Access Node must support the following scheduling schemes at each subscriber egress port: Strict priority and weighted mechanisms. Strict priority can be used for prioritising real-time traffic in case of wanting to assure that delay is kept to a minimum, whereas WFQ can be used for assuring that none of the traffic classes will suffer starvation.
- An Access Node must support policing per (upstream) traffic class at the subscribers ingress ports.
- An Access Node must support (re)marking of (upstream) traffic at the subscribers ingress ports.

### Requirements at the Access Node per network port

- An Access Node must support queuing/shaping per aggregated (upstream) traffic class at each network egress port.
- An Access Node must support the following scheduling schemes per network egress port: Strict priority and Weighted Fair Queuing (WFQ). Strict priority can be used for prioritising real-time traffic in case of wanting to assure that delay is kept to a minimum, whereas WFQ can be used for assuring that none of the traffic classes will suffer starvation.

### Requirements for configuration

- For each subscriber port the QoS Profile that defines the QoS values for that Port both in ingress and egress directions per Traffic Class (Parameters are e.g. SIR, PIR, Burst sizes) must be configurable by the management plane.
- For each network port, the QoS Profile that defines the QoS Values per aggregated Traffic Class for the egress traffic must be configurable. These profiles may be dynamically adjusted.
- The Mapping Table from .1p or DHCP Bits of ingress traffic to the supported Traffic Classes must be configurable (per subscriber Port). This mapping may be dynamically adjusted.

### Requirements for monitoring

- For each network port, reports on the usage of each traffic class will be generated and forwarded to the management plane upon request or once a predefined threshold has been reached.

### Requirements for multicast traffic

There is currently no standard way of triggering multicast resource requests. If an explicit request is to be sent from the end user terminal before the multicast process is started, an additional protocol would be needed and this would have a negative effect on zapping times. Another possibility would be to act on the receipt of an IGMP request for joining a specific multicast group, but IGMP has no mechanisms for including resource requirements or notifications of failed requests.

The following general requirements related to multicast CAC have been identified:

- A mechanism to trigger multicast resource requests.
- A mechanism to optimize multicast periodic requests.
- A mechanism to prevent the Access Node from starting multicast replication before the resource request completes and is granted.
- The ability to notify the client application about a failed multicast request.
- The ability to integrate admission control for unicast and multicast traffic.

## 5 CONCLUSIONS

MUSE advocates a pragmatic and simple way to provide services with QoS which is based mainly on traffic class differentiation, selective CAC and appropriate network dimensioning.

MUSE recommends a user-centric approach where classification of traffic into traffic classes is a responsibility of the user even if it is expected to be normally delegated to the providers. However, the network will ascertain that the usage of each traffic class does not exceed what has been planned, agreed or contracted by means of traffic policing. For the upstream traffic, it is recommended that the policy enforcement be done at the Access Nodes per Traffic Class and per Access Line. However, for the downstream traffic, it is recommended that the policy enforcement be done at the Edge Nodes per Traffic Class. It is also recommended to have per user shaping at the Access Node of downstream traffic to prevent congestion in the first mile.

The link between the Residential Gateway and the Access Node is usually the main bottleneck, especially since its capacity cannot easily be increased. Classification of upstream traffic into the traffic classes, and its prioritization onto the access link will be realized by the RGW according to user preferences. Additionally, traffic policing can be used to prevent starvation of lower priority upstream traffic and enforce the usage limitations per traffic class. Prioritization of downstream traffic per traffic class is made by the network according to the rules described in the user contract.

MUSE recommends the use of central CAC for the small subset of services that actually need it and only in those parts of the network where the network operator has identified a potential dimensioning problem. Local CAC at the Access Nodes or no CAC can be used for the rest of the traffic. Appropriate network dimensioning will help to minimize the risk of congestion or blocking problems.

Central CAC is recommended for applications requiring explicit CAC where neither scalability nor setup delays are special concerns (e.g. VoD). Local CAC is recommended to handle multicast traffic (e.g. IPTV) because of its better scalability and lower reaction time.

MUSE recommends a “provisioning” scenario where the central CAC, which has a view of all network resource usage, is able to allocate to a local CAC a certain amount of resources that will then be managed locally. Within this approach, the central CAC regularly monitors the usage of local resources at the Access Nodes and readjusts the resources allocated to the local CAC entities if necessary.

It is recommended that all network nodes support at least the four traffic classes proposed by MUSE (real-time, streaming, transactional and best-effort). At every link, outgoing traffic will be placed into a different queue according to what traffic class it belongs. These queues will be served by SP and/or WFQ scheduling mechanisms. SP is recommended to be used for guaranteeing a low delay for real time traffic in those links where capacity is relatively low, whereas WFQ can be used for the rest of classes. When capacity is high enough, WFQ can help to provision the required bandwidth to each traffic class.