Hierarchical QoS, The Next Step In Enhancing DOCSIS QoS Architecture

Alon Bernstein (alonb@cisco.com), Pawel Sowinski (psowinsk@cisco.com)
Cisco Systems, Inc.

Abstract

Robust and granular QoS support is a key factor in the tremendous success DOCSIS has achieved as a mature technology in delivery of data, voice and video services. The basic DOCSIS QoS architecture is built upon the foundation of the service flow concept that has been defined fourteen years ago as part of DOCSIS 1.1. The fundamental DOCSIS QoS framework has remained mostly unchanged since then.

Recently, as part of the industry dialog leading to CCAP and DOCSIS 3.1, cable operators have requested to apply QoS policies to aggregation of service flows in addition to individual service flow QoS. The paper presents a case for extending DOCSIS to include elements of hierarchical QoS (HQoS) technology for this purpose. While HQoS techniques have been deployed in other broadband access technologies, integration of HQoS with DOCSIS has never been attempted before. Such integration poses a number of unique technical and business challenges that deserve careful examination. Through the review of typical use case examples, the paper examines how the HQoS technology provides the business value to the cable operators in extending services to gigabit and beyond.

Finally, the paper explores a set of issues and potential solutions imperative to enable seamless integration of hierarchical QoS into CCAP deployments and to facilitate rudimentary multi-vendor interoperability, including necessary signaling protocol and elements of standard CMTS/CCAP configuration.

DISCLAIMER

The ideas described in this paper are part of Cisco's contribution to DOCSIS 3.1 specifications.

DOCSIS 3.1 specifications are still under development. The following represents authors' current thoughts on what HQoS in DOCSIS might look like, and do not represent actual decisions made regarding the final form of the specs or technology.

Anything could change.

INTRODUCTION

Brief history of DOCSIS QoS

In the late 90's cable systems were deployed without QoS. The assumption was that the 27Mbps over a 64QAM digital cable channel was more than enough bandwidth and when there is no congestion there is no need for QoS. It turns out that the above assumption was not true; TCP/IP which is the transport building block of the internet is "greedy" by nature (i.e. it attempts to fully utilize whatever pipe it has). On top of that a good percentage of cable subscribers are running greedy applications such as file sharing and servers.

QoS became a competitive issue; telecom companies came up with commercials claiming that a neighbor's activity can restrict a subscriber bandwidth. This was not a fair statement (the telecoms had their own aggregation bottleneck at the DSLAM output and in some cases those were less then 27mbps) but it did demonstrate the need to include QoS mechanisms in the cable access.

Since the need for QoS became obvious, the first version of DOCSIS (1.0) supported basic QoS. Each subscriber had a "QoS class of service" with traffic SLA (Service Level Agreement) defined per cable modem for the upstream and the downstream. While the above model helped resolve the congestion issues it clearly did not support multimedia service types for cable modems, e.g. voice/video/data.

The Requirement to support voice presented a clear and revenue-generating-reason to define QoS mechanisms that will assure voice quality even when the network becomes congested. In order to support voice and other multimedia definitions DOCSIS 1.1 changes the QoS model significantly:

1. The DOCSIS 1.0 class of service was obsoleted and instead DOCSIS 1.1 defines "unidirectional service flows" that could be separately created for the US and DS
2. New scheduling modes were defined for US to minimize latency for multimedia applications
3. The provisioned/authorized/admitted /activated states where defined for service flows

Note that the DOCSIS specifications define a behavior and not an implementation. In that spirit, this document describes how an HQoS should behave without defining how to implement it.

<u>Current State of DOCSIS QoS</u>

The DOCSIS 1.1 QoS model is still the base model for DOCSIS 2.0 and DOCSIS 3.0. The most significant recent update has been the aggregate QoS model for DPOE.

<u>DOCSIS QoS Applications and Hierarchical QoSs</u>

Hierarchical QoS allows an operator to define QoS policies on an aggregation of flows. Hierarchical QoS is defined as a strict tree structure where the physical interface capacity is typically the root (or "parent") node. The word "strict" means that for a given child node there can be one and only one parent.

Hierarchical QoS can be implemented by means of policing, shaping and/or scheduling. Some methods include rate limiting, rate shaping and/or marking packets for Weighted Random Early Discard (WRED). Some scheduling methods include prioritization and/or Weighted Fair Queuing (WFQ).

The current DOCSIS QoS definitions do not support aggregate QoS policies, but one can argue that the aggregation of all DOCSIS service flows into a single physical channel forms a simple two-level hierarchy where the physical level (in this case a QAM channel) is the parent node and the DOCSIS service flows are the children nodes. This is shown in the figure below.

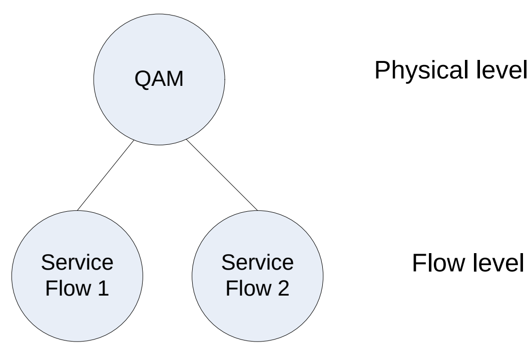

Figure 1: DOCSIS QoS model hierarchy

Hierarchy can be nested to create more complex trees. The paper will outline some of the use cases for these more complex hierarchies.

Hierarchical QoS and Fairness

Hierarchical QoS assures fairness across the children at a certain level, but not below or above. The following example can help explain the type of fairness that may be delivered in contrast to the DOCSIS 3.0 fairness model.

Fairness in the current, flat DOCSIS QoS model: if 1000 flows of equal priority share one congested interface then each flow gets 1/1000=0.1% of the bandwidth.

Fairness within Hierarchical QoS: Let's say we separate the traffic onto two virtual pipes, one for ISP A and one for ISP B, each limited to 50% of the total bandwidth, and use Hierarchical QoS:

If ISP A has only 1 active flow, and ISP B has the other 999 flows then:

- The single ISP A flow will get 1/2 = 50% the bandwidth
- Each ISP B flow will get (1/2)/999 = 0.05% of the bandwidth

In other words, Hierarchical QoS can assure fairness within the virtual pipes that it carves, but not across them.

HQoS in Edge Routers and Other Access Technologies

In the telecom world, hierarchical scheduling has been used for many years because of the multiple congestion points between the BRAS (Broadband Remote Aggregation Server) and the CPE (Customer Premise Device). Those include the interface to the DSLAM (Digital Subscriber Line Access Multiplexer) and the twisted pair to the CPE. Each one of these bottlenecks can be modeled as a "logical pipe" in a hierarchical scheduler. This use case was not needed by cable subscribers; however, as discussed in this paper, other use cases have emerged and renewed the interested in hierarchical QoS.

USE CASE REVIEW

In the initial discussions the cable operators have identified a number of use cases for HQoS. While those use cases differ in some details, such as the scaling numbers, the class of service (residential vs. business) or the types of service (data, video, and voice) in the end, the use cases boil down to two main service scenarios. These two scenarios are presented in detail below as subscriber level HQoS and service group level HQoS.

Use Case 1: Per Subscriber Aggregate QoS Controls

Recently, in order to effectively support more diverse service offerings the cable operators requested the ability to apply traffic controls, not only to individual service flows but also to groups of SFs belonging to a particular CM or IP host. The aggregate QoS limits must be enforced in addition to per Service Flow QoS treatment. The aggregate QoS enables the cable operators to offer SLAs with a simplified external structure and to more effectively compete with other ISPs which for many years have been providing internet access with hierarchically organized services.

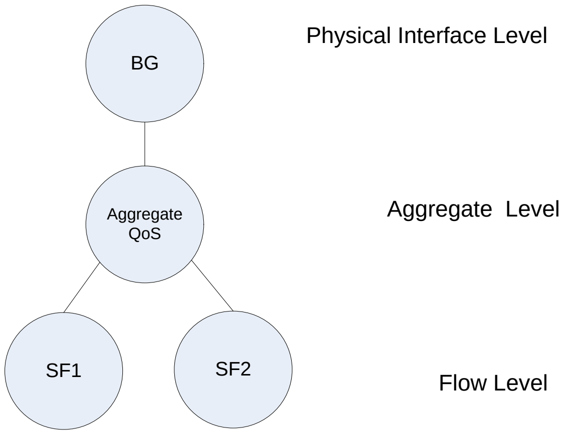

Figure 2: Use Case 1, Per Subscriber Aggregate QoS

As a typical example of use case 1, Figure **2** may be interpreted as a depiction of the QoS constructs for a downstream residential service with two QoS applications and a common, aggregate QoS policy. The individual applications are provisioned with Service Flows so that their traffic can have independent QoS treatment.

Let's consider the following settings:

- SF 1 represents a High Speed Data service with offered Maximum Sustained Traffic Rate at 20 Mb/s with Traffic Priority of 1.
- SF 2 represents a Managed Video service with offered Maximum Traffic Rate of 9 Mb/s with Traffic Priority of 5 as well as a Minimum Reserved Traffic Rate of 3 Mb/s.
- The aggregate QoS settings limit the combined traffic rate of both flows to 20 Mb/s.

The SLA which is structured with two levels of QoS controls gives the subscriber a single overall service rate at 20 Mb/s but allows the Managed Video service to operate with higher priority and with bandwidth reservation to guarantee a minimum level of QoS. When no traffic flows through the video service flow, the HSD service can use the whole 20 Mb/s. When the managed video is active, its traffic eats into the bandwidth offered to HSD service. Without aggregate QoS limit, both offered applications would run independently and potentially consume bandwidth at a higher cumulative level up to 29 Mb/s.

<u>Use Case 2: Virtual Partitioning of a Physical Interface or a Service Group</u>

As service offering diversifies, the cable operators would like to manage bandwidth allocated to each service from the total pool available in DOCSIS service group. Since the business model of HFC is largely relying on the concept of over-provisioning, during the periods of high usage (busy hours) the services compete for bandwidth and can negatively impact each other. This is where issues of fairness and the business model intersect. For example, a cumulative usage of bandwidth by a service with higher traffic priority can restrict bandwidth from service operating with lower DOCSIS traffic priority. It's imperative that the operators have effective tools to deal with such problems.

Today, DOCSIS 3.0 configuration offers the cable operators a convenient mechanism for this purpose. Operators may separate services by creating downstream bonding groups from distinct set of channels and appropriately "steer" the traffic belonging to each service. Such solution, based on partitioning of channel resources seems feasible considering that DOCSIS 3.0 DS channel pool in a SG scales into the tens and that each channel has relatively small capacity (38.8 Mb/s).

This method has its drawbacks as it is not applicable to upstream direction in HFC plants with low-split and mid-split. Additionally, physically partitioned service groups support lower peak traffic rates and don't share excess bandwidth. Further, with DOCSIS 3.1 such service separation scheme is no longer practical, as the channel bandwidth grows dramatically (1.7 Gb/s) and the channel count may drop down to just a few without enough granularity to effectively use them for traffic engineering.

PROPOSED DOCSIS QOS EXPANSION

Approach

When planning an expansion to an existing networking protocol, especially when it is deployed as widely as DOCSIS, one has to carefully consider a number of factors that define the overall fit of the newly added functionality into the existing architecture. These criteria, including backwards compatibility, consistency with the current methodology, multivendor interoperability as well as the ability to support incremental deployments have been contemplated when deciding the approach to introduce HQoS into DOCSIS.

In the end, we believe that the HQoS framework presented within this paper demonstrates a healthy compromise between the cable operators' requirements and how they fit into the existing DOCSIS architecture. The proposed framework includes the definition of devices' roles, key constructs, protocol signaling and common CMTS configuration.

On the other hand, the proposal does not dive into the details of implementation or internal algorithms used for queuing and scheduling. Further, the last component of the proposed framework, the CMTS configuration, may be extended in vendor-proprietary manner to meet individual vendor's needs and to provide solution differentiation.

The paper discusses HQoS primarily in the context of real-time traffic engineering and QoS policy enforcement. HQoS does impact non-real time functions such as admission control and resource management. These functions have to accommodate HQoS and the new policy controls HQoS provides. However, since these functions fall out of scope of current DOCSIS standards we feel there is no need to incorporate them in the proposed HQoS framework.

Lastly, it may be useful to note that the ideas describes in this proposal are symmetrical; they are equally applicable to traffic control for upstream and downstream directions.

The roles of CMTS and Cable Modems

HQoS is proposed as a CMTS only feature. The CMTS is responsible for all HQoS configuration and management. All aggregate QoS policy enforcement functions, including the real time traffic scheduling and queuing are performed only by the CMTS. The CMTS provides all network management capabilities necessary for status reporting related to HQoS. Cable Modems are not aware of the HQoS. CMs are required to convey HQoS information from CM configuration file into Registration Request without the need for interpretation of transported information. CMs need only implement certain QoS functions related to upstream bandwidth request policing on per SF basis only, as it is done in DOCSIS today.

New QOS Constructs

The key new construct introduced by HQoS is the Aggregate Traffic Class or the ATC. An ATC constitutes the middle point in the scheduling hierarchy; a point which is located in between service flows and physical interfaces. An ATC represents a group of service flows, or more precisely the aggregate of traffic flowing through a defined set of service flows.

As service flows, all ATCs are unidirectional; all service flows grouped into an ATC must serve the same upstream or downstream direction. In case of need for aggregate QoS policy enforcement for both directions, separate ATCs must be defined to group upstream flows and for a group of downstream flows.

Service flows may be mapped or associated with a single ATC through methods explained further in the paper. On the other hand, each ATC must be mapped onto exactly onto one physical interface to maintain hierarchical organization.

An ATC may group service flows with different QoS parameters. For example an ATC may aggregate traffic from service flows with different traffic priorities or different Maximum Peak Rates or maximum sustained rates.

ATCs have to be defined as part of the HQoS framework because they are externally visible; the information related to ATCs is exchanged in DOCSIS protocol and in the CMTS configuration.

Aggregate QoS Parameters

The CMTS enforces traffic control policy on an ATC. For this purpose, each ATC has an associated set of Aggregate QoS Parameters (AQPs). The parameters quantify the traffic control policy enforced by the CMTS. The set of proposed AQPs is listed below:

- **Aggregate Maximum Traffic Rate.** This is a mandatory parameter. It defines the maximum rate that the CMTS enforces on the aggregate of traffic flowing through the ATC. The choice of a specific algorithm for real-time enforcement of this parameter is left to vendor defined implementation.
- **Weight**. Weight controls arbitration when multiple ATCs “compete” for bandwidth of an interface. “Weight” should not be confused with “Traffic Priority”. Weight is an optional parameter.
- **Aggregate Minimum Reserved Traffic Rate.** By default, an ATC will operate with the Minimum Reserved Traffic Rate which is the sum of the values of Minimum Reserved Traffic Rate parameter for each of its member SFs. This parameter is provided to allow the operator to override the Minimum Reserved Rate value “inherited” from SFs. Aggregate Minimum Reserved Traffic Rate is an optional parameter.

Note that the list of proposed AQPs does not include all QoS parameters that can be defined for individual Service Flows. Some of the excluded parameters such as Traffic Priority or Maximum Burst size can be only defined at the Service Flow level.

Two methods are proposed to provision Aggregate QoS Parameters. The first method is based on explicit inclusion of an AQP set within an ATC definition. The second, indirect method relies on creation of named AQP profiles within the CMTS’s device configuration. This method resembles the existing DOCSIS mechanism for defining named Service Classes. Each AQP profile is identified by a name in the form of a string. The AQP profiles can be correlated to ATCs by name. Note, that when the AQP set is defined as part of the CMTS configuration the parameters can be augmented to include vendor proprietary extensions.

ATC Categories

This paper proposes two distinct categories of ATCs: Subscriber ATCs (SATCs) and Interface ATCs (IATCs). SATCs and IATCs differ in:

1) The purposes they fulfill
2) The selection and scaling of service flows mapped to them
3) The methods for provisioning and instantiation

Subscriber ATCs

Subscriber ATCs serve as an implementation tool for service layer agreements with two levels of QoS parameters. SATCs provide the outer QoS envelope, while service flows define QoS parameters for more granular, individual services or applications.

Each SATC aggregates traffic from a subset of unicast service flows associated with a single subscriber or a single Cable Modem. Not all service flows defined for a particular CM have to be a part of an SATC. Certain service flows may be directly mapped into physical interfaces. While Use Case 1 called for a single aggregate QoS per subscriber, the operator may be able to create more than one SATC per a cable modem. In such case, one set of CM's SFs may be mapped to one SATC while other SFs are mapped to other SATCs. As a general rule, a SF cannot be mapped to more than one ATC.

Figure **3** presents an example of a superset of Use Case 1 with five service flows and their relationship to two SATCs. In the example all five service flows belong to a single Cable Modem. SATC A includes service flows #1 and #2. SATC B consists of service flows #3 and #4. Service flow #5 is directly associated with bonding group BG Y so it is not mapped to any SATC.

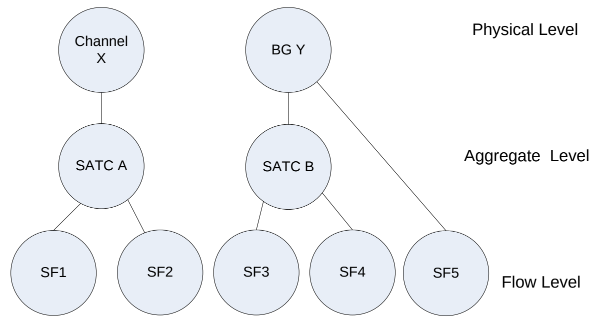

Figure 3: An example of SATCs for a single CM

How are SATCs configured? As a highly scalable edge platform, the CMTS does not maintain per subscriber configuration elements. Configuration objects that have per subscriber scaling, such as service flows, are typically configured via DOCSIS provisioning systems. Such objects are instantiated at the CMTS when a CM registers and conveys the content of its configuration file to the CMTS. SATC configuration and instantiation methods follow the practices devised for service flow provisioning. Such approach not only promotes scalability but also enables straightforward assignment of service flows to SATCs.

Operators provision SATCs by including their definition in the CM configuration file. The CM configuration file encodings are augmented to permit SATC definition. Table **1** lists the newly added CM configuration file encodings:

Attribute Name	Description
SATC Reference Number	A number identifying SATC in the CM configuration file
AQP Set or SATC AQP Profile Name	A set of scheduling parameters quantifying the QoS policy enforced by the SATC or a string which provides a reference to a named SATC AQP Profile. SATC AQP Profiles may be configured at the CMTS. The attributes: "AQP Set" and "SATC AQP Profile Name" are mutually exclusive as they provide alternative methods for provisioning of aggregate QoS policy parameters.
Service Flow Matching Method	A method by which dynamically created service flows can be matched to the SATC. The options for SF matching method are: 1. by SF Application Id 2. by SF priority range 3. by SF SCN
Service Flow Matching Criteria	A set of criteria which are dependent on the selected SF matching method.

Table 1: SATC Attributes

The service flows which are statically provisioned in the CM configuration file can be matched to SATCs by the SATC reference number as shown in

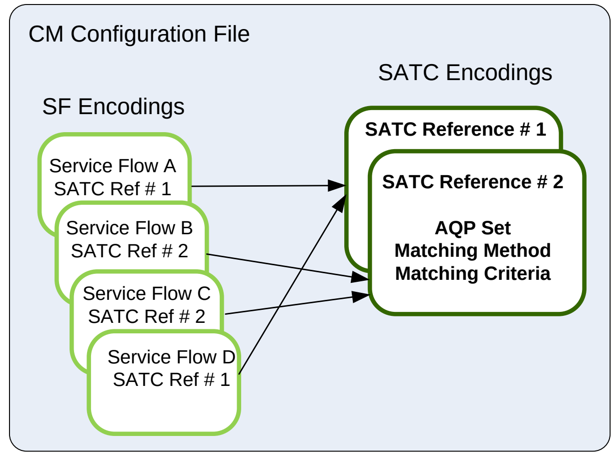

Figure **4**.

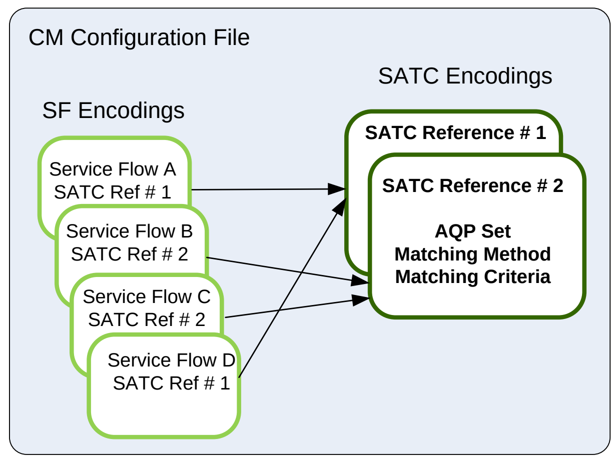

Figure 4: Correlation of SFs to SATCs in CM configuration file.

Dynamically provisioned service flows, for example those flows that are created through PCMM interface, may be matched to an SATC by one of the listed matching methods by means of the SF matching criteria.

Even though SATCs exhibit many similarities to service flows, the analogy between these constructs has limitations. Unlike DOCSIS Service Flows, SATCs don't maintain the QoS state attributes (provisioned, admitted, active), multiple QoS parameter envelopes or require associated QoS state management protocol. Cable Modems are generally not aware of SATCs.

Theoretically, the CMTS could be instrumented to create SATCs without explicit configuration; none of the identified use cases necessitates the dynamic creation of SATCs. The need for SATC QoS state management is further abated because the CMTS's admission control functions must operate at Service Flow level. Therefore the paper asserts that SATCs should be generally considered to be static objects, always present and active after a CM completes its registration.

Recently, DPoE specifications introduced into DOCSIS the concept of Aggregate Service Flows (ASFs). ASFs have been added to DOCSIS proper (MULPI) for the purpose of reserving TLV numbers. ASFs in DPoE and SATCs as proposed here for DOCSIS provide largely equivalent functionality and are provisioned in a similar way. They differ in the environment for which they have been designed and certain operational requirements. For example ASFs in DPoE may be associated with a full Service Flow QoS Parameter Set, in some cases these parameters are overridden by equivalent MEF parameters. The AQP proposed for SATC includes fewer parameters. While the authors believe that these concepts of SATC and ASF can be merged, doing so is outside of the scope of the paper.

<u>Interface ATCs</u>

Interface ATCs are intended to fulfill the premise of the Use Case 2. The IATCs enable the operators to virtually divide the bandwidth of service groups or physical interfaces between distinct services or users.

An example of such partitioning is shown on

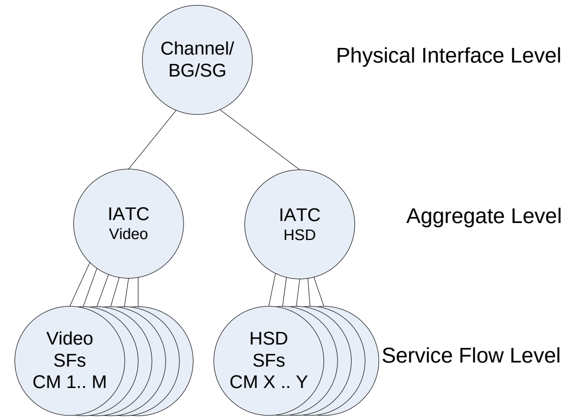

Figure **5**.

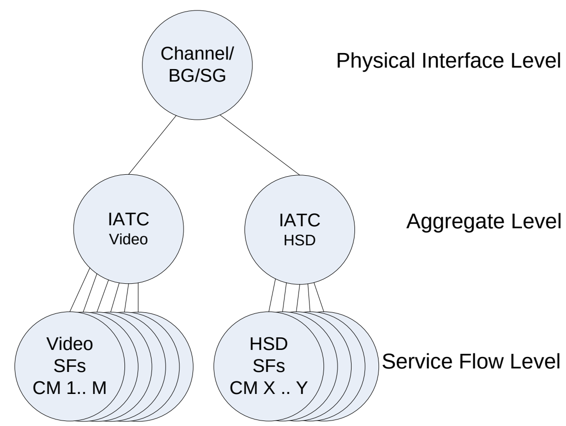

Figure 5: Interface ATCs

An IATC aggregates traffic from Service Flows belonging to multiple CMs but typically sharing some common property like application, the type of service or selected traffic priority. For example, one IATC may be created to group traffic from all service flows carrying cable operator's managed video traffic that are mapped to a particular bonding group. In the same example a second IATC may group all HSD flows mapped to the same bonding group. The aggregate QoS parameters associated with each IATC will define how the bandwidth of the bonding group is shared between managed video and HSD.

IATCs are provisioned via the CMTS configuration in a two-step process. In step one; the operator defines a number of IATC profiles. IATC profiles are identified by a name and can be used throughout the system. IATC profiles serve as templates for creation of IATC instances. Each IATC profile includes the attributes as listed in Table **2**.

Attributes	Description
IATC Profile Name	A string that uniquely identifies the IATC profile.
Aggregate QoS Set	A set of parameters defining the QoS policy enforced by the IATC.
SF Matching Method	A method by which the CMTS can match service flows (both static and dynamic) to the IATC. The following methods are proposed for SF matching: • by Application Id • by SF priority range • by SF SCN • None Note: "None" matching method may be selected when statically defined service flows in CM configuration file are explicitly matched to an IATC profile by name.
SF Matching Criteria	The set of criteria that corresponds to the configured matching method: Application Id, SCN, SF traffic priority range.

Table 2: IATC attributes

In step two of the configuration process an operator can associate any selected physical interface with one or more IATC Profiles. When more than one IATC profile is associated with an interface then the SF matching method or SF matching criteria must differ between IATC Profiles to ensure unambiguous matching decision. Not all physical interfaces must be paired to an IATC Profile. Figure **1** demonstrates an example of configuration defining the association between static physical interfaces and IATC profiles.

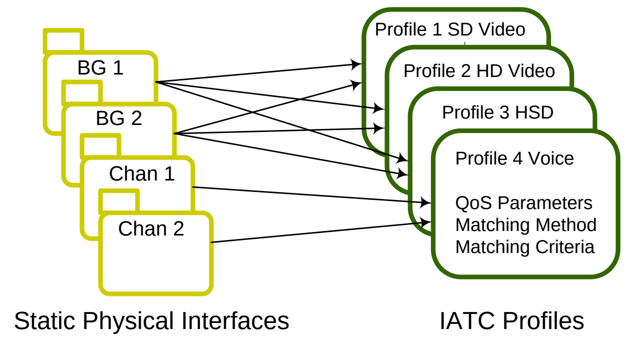

Figure 6: Static Physical Interfaces to IATC Profiles Mapping

Based on such configuration, the CMTS creates instances of IATCs for each configured pair {interface, IATC profile}. In the example shown on Figure 10, 3 IATC profiles will be instantiated for each bonding group, BG 1 and BG 2 and one IATC profile will be instantiated for each of the channels. All together the shown configuration results in creation of 8 IATCs.

The IATC provisioning method described above can be deployed for those physical interfaces that are created statically. DOCSIS allows CMTS's support for the dynamic creation of upstream or downstream bonding groups. Yet, this function is largely left to CMTS vendor definition because DOCSIS does not define a specific method or standard configuration for this purpose. This proposal takes a similar approach to the definition of HQoS over dynamically created bonding groups. We acknowledge that dynamic BGs can be associated with IATCs, but the specification of such method is left to CMTS vendor differentiation.

How are Service Flows mapped to IATCs? In the absence of H-QoS the CMTS maps Service Flows to bonding groups or individual channels. With HQoS the SF mapping process needs include one additional step: a decision whether to assign a SF to an IATC and which IATC to select. Operators will be able to control SF to IATC association via several matching methods. Those methods are defined as part of IATC configuration and listed in Table **2**.

An alternative mechanism permits association of SFs provisioned via CM configuration file to IATC profiles by name. The SF encodings in the CM configuration file are augmented with IATC name for this purpose.

IATCs typically group a subset of service flows from a service group, scaling up to hundreds of SFs. SATCs will typically aggregate traffic from a much smaller SF grouping, perhaps reaching into the teens.

The Impact on DOCSIS Protocol

The presented HQoS framework has a small impact on the current DOCSIS protocol. The scope of changes is restricted to augmentation of the CM configuration file encodings to support a few added HQoS TLVs. The CM opaquely conveys the new TLVs to the CMTS during registration. Otherwise, HQoS does not require any additional support by the CM software or hardware.

CONCLUSION

DOCSIS has been a very successful protocol. The granular QoS support in DOCSIS is part of this success, and can continue as such especially if its functions can be adapted to better serve the new range of services. This paper proposes a natural expansion to DOCSIS QoS to provide the cable operators with control over QoS policies at the aggregate level. The expansion creates tools to enable the operators to offer new and better structured services to their customers as well as to more effectively manage the allocation of one a most valuable resources at their disposal: the bandwidth of the DOCSIS part of the HFC.

ACKNOWLEDGMENTS

We would like to extend our sincere gratitude to many talented participants of the DOCSIS 3.1 MAC Focus Group and the AMP Work Group for their comments and contributions to the paper.

List of Acronyms

AQP – Aggregate QoS Parameters
ASF – Aggregated Service Flow
ATC – Aggregate Traffic Class
HQoS – Hierarchical QoS
HSD – High Speed Data
IATC – Interface ATC
MEF – Metro Ethernet Forum
QoS – Quality of Service
SCN – Service Class Name
SF – Service Flow
SG – Service Group
SATC – Subscriber ATC

References

1. DOCSIS MULPI : http://www.cablelabs.com/specifications/CM-SP-MULPIv3.0-I20-121113.pdf
2. Packet Cable MultiMedia Specification: http://www.cablelabs.com/specifications/PKT-SP-MM-I06-110629.pdf
3. AMP Technical Report: MAC Layer Design for Efficient Multi-Gigabit Transport over CATV Networks. CableLabs.