A Concept of inter-AS Priority Signaling using BGP Attributes

Th. M. Knoll, Chair of Communication Networks, Chemnitz University of Technology, Germany
Email: knoll@etit.tu-chemnitz.de

Abstract—The increasing number of delay and loss critical services in packet networks require differentiated packet handling in the forwarding plane. Quality of Service (QoS) guarantees can be given for networks using resource reservation and admission control. However, such strategies require complex control plane extensions and might lead to higher operation expenditures.

Network operators therefore often use over-provisioning and traffic differentiation to offer cheaper class of service quality in their internet protocol (IP) packet networks.

Priority marking and forwarding of packetized data traffic can be realized mainly using different layer two and three mechanisms. The number of differentiated classes and their autonomous system (AS) internal implementation is at the operator’s choice.

This paper proposes a concept of cross-domain and cross-layer priority signaling between packet switched networks to be used at the inter-AS peering points. It makes use of the Border Gateway Protocol (BGP) as the predominantly used routing protocol for AS peering communication. A new BGP-4 path attribute is used to convey the structured priority information.

The new concept ensures that all receiving AS peers are consistently and comprehensively informed about the QoS handling within the IP prefix originating AS. Based on this information, all ASes can perform close QoS treatment approximation in a cross-domain and cross-layer manner.

The approach is now work in progress at the IETF.

Index Terms—BGP, cross-domain priority signaling, inter-AS QoS, QoS Class extended community attribute

I. INTRODUCTION

Quality of Service (QoS) can be achieved using either resource reservation with admission control or through service differentiation based on prioritized traffic classes. This paper focuses on traffic class priority only and will support coarse QoS in terms of “Class of Service (CoS)”. For that, different layer 2 (L2) and layer 3 (L3) priority mechanisms are used today. Autonomous System (AS) providers can freely choose and configure the set of mechanisms to be used within their network. These AS-internal QoS policy decisions are made independently and will not necessarily be shared or synchronized with neighboring ASes. This paper proposes a modified Border Gateway Protocol (BGP – see section IV) QoS signaling mechanism, which provides a consistent unidirectional cross-domain QoS information exchange.

AS-internal mapping between priority mechanisms of different network layers is not standardized either. However, this new signaling mechanism will signal the applied mapping within the IP route originating AS to all other ASes. Network operators can use this originating AS cross-layer mapping information to adopt their local QoS mapping and packet treatment on all QoS enabled network layers. QoS in this approach refers to primitive traffic separation into several classes, which will experience differently prioritized forwarding behavior in relaying nodes. No QoS parameters are guaranteed, but enqueuing in separate forwarding queues is aspired, which leads to a better than best effort forwarding behavior. The approach is currently implemented for experimental analysis in the so called Quagga routing suite under the Linux operating system.

This paper first outlines a general overview of existing layer 2, layer 2.5 and layer 3 priority mechanisms and their respective QoS classes. It highlights the difficulties in cross-domain as well as cross-layer QoS signaling and mapping in section III. The major contribution of this work is described in section IV. BGP is used as signaling transport mechanism for a new BGP QoS Marking Attribute. Section IV gives a short overview on BGP operation and the BGP message structure before the actual attribute definition follows.

The last five sections explain the proposed signaling on an example network, outline its usage, name related work and consider security, confidentiality and business aspects.

II. OVERVIEW OF COMMONLY USED PRIORITY MECHANISMS

Table I gives an overview of the number of supported classes of service by different link layer and network layer technologies.
The number of available classes and the respective marking mechanism influence the forwarding treatment granularity and observability. Several connection-oriented technologies allow for precise quality specification through a set of traffic parameter values. This paper, however, focuses on relative “class of service” categories and will therefore assume a mapping of such parameter sets into major QoS categories. The following three paragraphs will briefly explain layer 2, layer 3 and layer 2.5 mechanisms for priority based traffic differentiation.

A. Layer 2 Priority

Data Link Layer technologies provide different priority mechanisms and markings. Such QoS enabled layer 2 technologies can adopt their queuing and scheduling behaviour according to the marked priority of each frame. Connection-oriented technologies can additionally provide separate forwarding paths (tunnels) for different priorities.

1) Connectionless technologies (e.g. Ethernet)

Ethernet is the predominant connectionless layer 2 technology and provides 8 possible priority classes. So called “tagged frames” are required in order to accommodate the necessary marking bits. The marking is done by 3 bits called “user priority” within the priority tagged frame structure as defined in IEEE 802.1q [15]

This standard also describes recommended user priority to traffic class mappings depending on the targeted number of differentiated classes. Furthermore, it describes outbound access priorities when interfacing with Wireless LAN, Token Ring, FDDI and other layer 2 technologies.

Interestingly, there are no precise definitions as to which treatment is associated with a specific priority marking. Furthermore, no priority mappings between 802.1q and layer 3 technologies are provided.

2) Connection-oriented technologies (e.g. UMTS, ATM, FR)

All connection-oriented layer 2 technologies provide separate virtual channels between endpoints. Each channel – represented by a certain channel ID (ATM VPI/VCI, Frame Relay DLCI etc.) – is usually associated with certain traffic parameters, which are used for service differentiation and resource reservation. The traffic parameters are generally too detailed and complex for the coarse traffic class differentiation focused on in this paper.

However, those connection-oriented technologies still distinguish major traffic classes. For example the Universal Mobile Telecommunications System (UMTS) [1] defines four traffic classes for the packet switched access mode (“Conversational class”, “Streaming class”, “Interactive class” and “Background class”). Asynchronous Transfer Mode (ATM) also provides four major traffic classes, called “CBR”, “VBR”, “ABR” and “UBR”. Frame Relay does not define traffic classes, but providers may setup several virtual channels and associate different resource reservations with them.

B. Layer 3 Priority

The only network layer protocol discussed in this paper is the internet protocol (IP). Both common versions, IPv4 and IPv6, provide bits in its packet header structure for priority marking. The original “Type of Service” and “IP precedence” encoding has been redefined within the Differentiated Services (DiffServ) architecture [5]. The priority marking is realized using 6 bit Differentiated Services Code Points (DSCP) in place of the IPv4 - TOS/IP precedence as well as the IPv6 - Traffic Class octet [21]. The remaining two bits are not used for priority marking.

This 6 bit field allows in theory for the marking of 64 different traffic classes. The DiffServ architecture defines about 15 so called “Per Hop Behavior (PHB)” and their associated DSCP values. The number of locally applied PHBs and the respective DSCP encoding can be chosen by network operators following local policies.

This overwhelming freedom of marking and associated forwarding treatment is targeted by RFC4594 [3], which gives “Guidelines for DiffServ Service Classes” for network administrators. Depending on the planned number of differentiated classes, clear recommendations on the DiffServ class selection, the DSCP encoding and the forwarding treatment mechanisms are made.

C. Layer 2.5 (MPLS) Priority

Multiprotocol Label Switching (MPLS) is widely used in carrier networks and provides two ways of traffic class differentiation.

The MPLS framing structure in Fig. 1 includes 3 “experimental” bits (EXP), which are commonly used for traffic class differentiation within established label switched paths (LSPs). This way, eight classes can be encoded which can be used to influence the packet treatment along the forwarding path. LSPs, which use EXP bit inferred forwarding treatment, are called “E-LSPs”. 

<table>
<thead>
<tr>
<th>Network type</th>
<th>Supported QoS classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP supporting DiffServ</td>
<td>64 (currently 21 defined)</td>
</tr>
<tr>
<td>IP supporting ITU Y.1541</td>
<td>6</td>
</tr>
<tr>
<td>Ethernet (IEEE 802.3)</td>
<td>8 (802.1q priority tag)</td>
</tr>
<tr>
<td>MPLS</td>
<td>8 (E-LSP) or 2^20 (L-LSP)</td>
</tr>
<tr>
<td>ATM</td>
<td>4 major QoS categories &amp; 2 dropping priorities each</td>
</tr>
<tr>
<td>UMTS</td>
<td>4 major QoS categories</td>
</tr>
</tbody>
</table>
III. CROSS-DOMAIN AND CROSS-LAYER QoS

Two different problems are identified for consistent end-to-end QoS handling. They are cross-domain QoS signaling and cross-layer traffic class mapping.

A. Cross-Domain QoS signaling

Differentiated end-to-end QoS is ideally achieved by identical QoS marking and packet treatment policies across all networks involved in the end-to-end packet transport. However, the local policy enforcement and possible remarking options will likely lead to uncoordinated sub-optimal packet treatment for traffic travelling through several ASes.

On layer 3, neighboring ASes are able to recognize the “original” DSCP encoding before the possible remarking to local values occurs. Other networks along the path, however, will not be able to refer back to the source behaviour encoding and adopt their (re-)marking accordingly.

There is no inter-AS routing protocol inherent mechanism known, which provides a signaling mechanism to remedy this limitation. Furthermore, layer 2 and layer 1 QoS mechanisms are not visible to neighboring networks at all.

Providing and enforcing guaranteed end-to-end QoS requires a complex control and signaling framework, which has been worked on in several projects before (see section IX).

This approach is not providing guaranteed end-to-end QoS, but makes the available traffic separations, encodings and cross-layer mapping visible to all ASes. It therefore enables route selection and marking adoption without guarantees and paves the way for transparent customer traffic transport by inter-domain tunneling with QoS based traffic separation.

B. Cross-Layer QoS mapping

Service class differentiation in one networking layer should ideally match the same differentiation in the upper and lower layers as well.

Network operators chose priority mechanisms on different layers depending on their local QoS policy and the available networking technologies. As mentioned before, there are no Cross-layer QoS mappings standardized between e.g. Ethernet user priority marking and IP DSCP marking. Such mapping occurs AS-internally in an uncoordinated manner and is not signaled across network boundaries.

C. Targeted QoS signaling strategy

Given the mentioned problems in consistent QoS signaling and mapping, this paper proposes a new QoS signaling strategy for cross-domain priority signaling and cross-layer priority mapping.

It is unlikely to synchronize all AS operators to a common set of QoS classes, encodings and packet treatment on all possible network layers. However, if network operators are informed about the applied QoS mechanisms and markings in neighboring networks, they can adopt their local transit packet handling to achieve a close QoS treatment approximation.

The first and currently used solution for this inter-AS information exchange can be based on mutual agreements (e.g. SLA agreements) about mapping and treatment procedures for transit traffic. However, if several transit ASes are involved in an end-to-end traffic path, several such mutual agreements are in force, which will most certainly lead to a falsified QoS treatment along the way.

The second solution for this inter-AS QoS information exchange is provided by this paper through a new inter-AS QoS priority signaling mechanism. The IP prefix originating AS performs the decision on QoS class selection, encoding and packet treatment on all networking layers based on its local QoS policy. The new mechanism will augment inter-AS routing UPDATE messages with that information about the locally applied QoS classes and their AS-internal networking layer dependent encoding. Furthermore, a grouping by the respective QoS class is performed.

This source oriented QoS transitive signaling will be transferred unchanged to all internetworked ASes.

The new strategy ensures that all receiving AS peers are consistently and comprehensively informed about the QoS handling within the IP prefix originating AS. Based on this information, all ASes can perform close QoS treatment approximation in a cross-domain and cross-layer manner.

IV. MODIFIED BGP QoS SIGNALING

The Border Gateway Protocol version 4 (BGP-4) is the predominant inter-AS routing protocol. It is the base for the proposed new QoS signaling mechanism of this paper and will therefore be explained in more detail in the following paragraphs.
A. inter-AS routing and signaling using BGP

IP networking between networks of different carriers is realized by peering points and associated peering agreements called “Service Level Agreement (SLA)”.

The BGP-4 is used for the mutual information exchange about reachable IP networks.

Fig. 2 BGP peering between neighboring ASes

Fig. 2 shows a simple network setup where AS border routers establish BGP peering sessions between configured neighboring peers and send BGP messages for a secured and reliable global routing information exchange.

The virtually fully meshed BGP peer overlay network allows for a controlled inter-AS information exchange, which can be used to signal reachability and other information between all existing AS border routers in a unique and consistent manner.

The core BGP routing information messages are so called BGP UPDATE messages, which consist of four parts:

1. Message header,
2. Withdrawn routes,
3. Path attributes and
4. Network Layer Reachability Information – NLRI.

This message structure can also be seen in Fig. 5.

The UPDATE messages include BGP Path Attributes, which signal origin and routing details. This information can be used to further control the route filtering and the advertisement process.

These attributes are of central concern for this paper and are therefore explained in detail.

BGP Path Attributes

A number of BGP path attributes are defined and can be grouped into “well-known vs. optional” (see Fig. 3) and “transitive vs. non-transitive” attributes.

B. Definition of the QoS Marking Attribute

A new extended community attribute is defined here, which carries QoS marking information for different network layer technologies across ASes.

This new BGP extended community attribute is called “QoS Marking Attribute” (Fig. 4). This new attribute provides a mechanism for labeling priority class information carried in BGP-4. It allows for the consistent exchange of priority class encoding values between BGP peers for physical, data link, network and transport layer QoS mechanisms.

These labels can be used to control the distribution of this information, for the encoding and for treatment adjustments within the AS or for other applications.

The QoS Marking Attribute is a transitive optional BGP attribute, with the Type Code 16. [12]

IANA has assigned the type code 0x00 for this Extended Community Attribute.

Several QoS Marking Attributes may be included in a single BGP UPDATE message.

Each QoS Marking Extended Community is encoded as 8-octet tuple, as follows (see Fig. 4):

- Flags : 1 octet,
- QoS Set Number: 1 octet
- Technology Type: 1 octets,
- QoS marking O: 2 octet,
- QoS marking A: 1 octet,
- Processing Count: 1 octet.

Fig. 3 BGP path attributes [10], [12], [23] and [24]

This paper focuses on the use of the so called “Extended Community Attributes”, which are defined in RFC4360 [24].

Community attributes (including extended community attributes) are not essential for BGP peering and proper interworking of inter-AS IP routing. They are optional attributes and can be marked as transitive (signaled across all ASes) or non-transitive (signaled only between neighboring AS peers). Community values are widely used today. They can for instance reference SLA defined mechanisms and route filtering rules between neighboring peers without global significance.

Fig. 4 QoS Marking Attribute structure
Flags:
All used and unused flags default to a value of ‘0’.
Table II shows the bit encoding of the flags field.

**Table II QoS Marking Attribute – Flags field**

<table>
<thead>
<tr>
<th>Bit</th>
<th>Flag</th>
<th>Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 2</td>
<td>unused</td>
<td>Default to ‘0’</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>‘1’ … remarking occurred</td>
</tr>
<tr>
<td>4</td>
<td>I</td>
<td>‘1’ … QoS marking ignored</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>‘1’ … QoS class aggregation occurred</td>
</tr>
<tr>
<td>6 &amp; 7</td>
<td>unused</td>
<td>Default to ‘0’</td>
</tr>
</tbody>
</table>

QoS Set Number:
Several single QoS Marking Attributes can be logically grouped into a QoS Marking Attribute Set characterized by a identical QoS Set Number. This grouping of the single QoS Marking Attributes into a set provides cross-layer linking between the QoS class encodings. It can also be used for the specification of behavior sets are given in the [RFC3140]. The number of signaled QoS Marking Attributes and QoS Marking Attribute Sets is at the operator’s choice of the originating AS. The enumerated QoS set numbers have BGP UPDATE message local significance starting with set number 0x00.

Technology Type:
The technology type encoding uses the enumeration list in Table III.

**Table III Technology Type Enumeration**

<table>
<thead>
<tr>
<th>Value</th>
<th>Technology Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00</td>
<td>DiffServ enabled IP (DSCP encoding)</td>
</tr>
<tr>
<td>0x01</td>
<td>Ethernet using 802.1q priority tag</td>
</tr>
<tr>
<td>0x02</td>
<td>MPLS using E-LSP</td>
</tr>
<tr>
<td>0x03</td>
<td>Virtual Channel (VC) encoding using separate channels for QoS forwarding / one channel per class (e.g. ATM VCs, FR VCs, MPLS L-LSPs)</td>
</tr>
<tr>
<td>0x04</td>
<td>GMPLS - time slot encoding</td>
</tr>
<tr>
<td>0x05</td>
<td>GMPLS – lambda encoding</td>
</tr>
<tr>
<td>0x06</td>
<td>GMPLS – fiber encoding</td>
</tr>
</tbody>
</table>

QoS Marking / Class Number O:
The interpretation of this field depends on the selected layer and technology. QoS mechanisms using bit encodings for the targeted priority (e.g. PHB ID Codes, Ethernet User Priority, MPLS EXP etc.) must use the same encoding in this attribute field. Inter-domain DiffServ behavior signaling uses so called PHB ID codes as defined in RFC3140 [4]. Unused higher order bits default to ‘0’.
Other technologies, which use other ways of labeling priority information such as L-LSPs, VPI/VCI inferred ATM classes, lambda inferred priority, etc. shall use class enumerations as encoding. This “O” field is only be set by the IP prefix originating AS. All relaying ASes are not allowed to change this value.

QoS Marking / Class Number A:
The “A” field is used locally to provider peering points. It conveys the actually applied markings at the peering point and can be changed by relaying ASes in order to adopt to the locally available QoS traffic separation. Layer 3 behavior signaling is hereby using the locally applied DSCP encoding.

Processing Count:
Each BGP peer, which adds a new AS number into the BGP AS_PATH attribute must increase the processing count in order to signal its processing of the marking attribute. Relaying ASes can thereby compare this count with the number of different ASes in the path. Identical counts will signal full support of the marking attribute along the forwarding chain of ASes.

The QoS Marking Extended Attributes are used in BGP UPDATE messages (see Fig. 5). Only the prefix originating AS is allowed to signal the QoS Marking Attributes and sets. Transit ASes must not modify or extend the class set except for the adoption of the Marking A and Processing Count field.
V. QoS MARKING ATTRIBUTE EXAMPLE

The QoS Marking Attribute provides QoS labeling for different priority encodings and technologies. Fig. 6 shows a simple example network setup with three peered ASes and the exchange of two QoS enriched BGP UPDATE messages. AS1 and AS3 use three priority classes on layer 3 and layer 2. The inner AS2, however, distinguishes only two priorities combined with QoS-based routing.

The example shows how DiffServ DSCP values, Ethernet user priority values as well as two MPLS E-LSP encodings are represented in the proposed QoS Marking Attribute.

The structure of the exchanged BGP UPDATE messages is similar to the structure in Fig. 5. The number of included QoS Marking Attributes varies according to the number of classes defined within the respective AS as well as the number of QoS supporting layers within the configured priority policy. BGP UPDATE messages with identical QoS Marking Attributes can be merged in relaying peers before being advertised to neighbouring peers.

Each AS in the example is represented as a yellow circle with its red BGP enabled border routers. The blue smaller sized nodes represent internal routers, which are IGP controlled and differently meshed with each other. This internal physical network structure is out of scope of this paper and not relevant to the proposed extended BGP functionality.

The QoS Marking Attribute part of two BGP UPDATE messages from AS1 to AS2 and AS2 to AS3, respectively are shown in Fig. 5. Differently colored packets are relayed from a not shown source left of AS3 and travelling towards a not explicitly drawn end system within AS1. The colors represent different priority markings. AS1 and AS3 obviously use QoS-based treatment for node internal forwarding treatment only. AS2, however, uses different MPLS tunnels (E-LSPs) for QoS-based forwarding and routing. In turn, only two classes are distinguished, which requires some sort of mapping between external QoS class sets into the AS2 specific QoS class set. This configuration is the AS2 operator’s choice and out of scope for this paper.
Given the two UPDATE message attribute sets, AS2 and AS3 are signaled, which QoS mechanisms and encodings are used by AS1 and AS2 for their originating IP prefixes. This information exchange between BGP peers occurs omnidirectional and between all peering BGP border routers of the Internet. This way, all ASes get an overview knowledge about the applied QoS policies in QoS Marking Attribute enabled ASes.

The following paragraphs discuss the resulting forwarding treatment and routing of IP packet traffic targeting AS1 and originating from AS1.

A. Sending packets towards AS1 (Traffic Target)

AS3 and AS2 know the QoS class set (mechanisms and encoding) from the traffic originating AS through its signaled “QoS marked” prefixes. For the packet transfer, this QoS class set is inferred from the packet’s source address. The actual class of the packet and the targeted treatment is inferred from the locally applied marking value on each networking layer.

The mapping of the original DiffServ, Ethernet, ATM VC, Frame Relay DLCI and so on marking mechanism into the
available mechanisms in the downstream AS neighbor is done at the traffic egress router with locally defined mapping strategies. Close approximation of the locally available QoS treatment to the neighboring QoS treatment is aspired. This assumption can be guided through SLA arrangements or more likely inferred from the combination of signaled mechanisms and marking values.

Transit ASes, which are not aware of the QoS Marking Attribute or which are not capable to perform an appropriate mapping must mark the respective attribute with the ‘P’ flag [23]. This activated partial bit signals other BGP peers that this attribute is not consistently applied to all transit AS networks. On the other hand, a cleared ‘P’ flag assures that all transited ASes will have interpreted the signaled attribute.

However, it can not be derived from this single flag alone that those transit networks will all perform QoS based traffic forwarding treatment towards the associated IP prefixes. The ‘R’ (“remarking occurred”) flag and the ‘I’ (“QoS marking ignored”) flag within the proposed QoS Marking Attribute signal the actual QoS efforts along the path.

B. Receiving packets from AS1 (Traffic Origin)

AS3 and AS2 know AS1’s QoS class set through the BGP UPDATE message for “QoS marked” AS1 originating prefixes. This QoS class set is inferred from the packet’s source address (e.g. 134.109.4.105). The actual class of the packet and the targeted treatment is inferred from the locally visible marking on the different networking layers. The flag marking and forwarding treatment efforts remain the same as described above.

VI. USAGE OF THE BGP QoS MARKING ATTRIBUTE

Providers may choose to analyze the QoS Marking Attributes and adopt the priority encoding according to their local policy. Whether this may also lead to different IGP routing decisions or even effect BGP update filters is out of scope for this paper.

Besides the layer 3 encoding and routing adaptation, the AS local mapping onto layer 2, 1 and 4 QoS mechanisms opens a new quality in prioritized cross-domain and cross-layer QoS mechanisms. This new quality in prioritized cross-domain and cross-layer QoS mechanisms opens a scope for this paper.

Direct mapping between e.g. IP DSCP and tagged Ethernet is not standardized in either standard. However, by means of the QoS Marking Attribute, all ASes – neighboring and non-adjacent ASes – are now able to refer to the originating AS encoding on either layer. This information allows for close to ideal service differentiation along the end-to-end packet forwarding path.

This signaling between AS border routers circumvents the difficulties associated with route and traffic class aggregation.

Current standardization efforts focus on layer 3 traffic class aggregation where the behavior encodings of aggregated packet streams will have to be mapped onto the locally selected common traffic class encoding. Such scenarios are discussed in RFC5127 [9]. IP prefix aggregation will be performed unchanged as today. The aggregating AS becomes the prefix originating AS for that prefix and applies its Class set onto it. However, it should keep track of the originally available class sets and provide internal mapping for such traffic. Aggregation must be signaled using the “A” flag in the attribute.

Tunneling of transit traffic (e.g. IP in IP encapsulation) is one possible solution to preserve the original encoding. It is a complex mechanism and creates at least 20 bytes of overhead for the preservation of a 6 bit encoding.

The idea of reduced class sets in core networks and restored encoding for transit traffic at network egress routers holds operational and scaling potential. AS-internal transparent tunneling of customer traffic is strongly recommended for this approach.

The same mechanism can be used e.g. in carrier grade Ethernet scenarios where Ethernet user priority regeneration is performed as described in 802.1q [15].

The proposed inter-AS priority signaling using BGP attributes provides the means to signal the locally applied mapping outside the originating AS. Transit ASes can therefore refer back to the original QoS treatment and marking strategies and adopt their local packet treatment accordingly.

Close QoS treatment approximation across ASes and across several networking layers within the forwarding path can be achieved using the QoS Marking BGP attributes.

VII. IMPACT ON ROUTING AND FORWARDING

BGP route updates using the proposed attribute will be able to signal different layer QoS encodings between AS boundaries (eBGP) as well as within ASes (iBGP). Network operators are able to adjust their internal marking and route advertisement strategy as well as traffic engineering based on the signaled information. This internal strategy can transparently being reverted at the AS egresses to the original source AS marking using the community attribute information.

This mechanism ensures consistent – update source based – QoS signaling as well as close treatment adaptation despite possibly different AS internal QoS policies. AS internal redistribution of QoS information using various reservation and prioritization schemes is not focused on in this paper.

The further usage of the conveyed QoS information could be twofold. Current BGP practice performs best path selection only within each BGP peer before advertising its path update information. That is, the information could be incorporated as one additional step in the best path selection algorithm. This mono-path approach would ensure, that the selected path does provide traffic separation along the forwarding path. BGP multipath would allow for QoS-based routing using class-specific path selection. However, this approach is hardly used in practice and will not be available globally in the foreseeable future.

One major concern for the BGP operation is the size of the
routing tables. About 250000 IP prefixes are currently stored and looked up in this routing information base, which places a strong burden on memory and processing power for BGP peers. The proposed aggregation procedure is therefore vital for BGP scalability and the acceptance of this approach.

Signaling QoS information does not lead to significantly increased routing update traffic. All route prefixes within a BGP UPDATE message are associated with the included QoS Marking Attribute Sets of the message.

VIII. IMPACT ON SECURITY AND CONFIDENTIALITY

The proposed QoS Marking Attribute does not raise extra security concerns. Existing BGP security measures are in place through the reused transport in BGP UPDATE messages.

The disclosure of confidential network intrinsic information is of no concern since network operators have full control on which encoding is signaled and for which IP prefixes.

Furthermore, if the network internal QoS mechanisms and markings shall not be disclosed to BGP peers, but the proposed BGP QoS signaling shall still be supported, it is also feasible to signal a “faked” set of QoS mechanisms and encodings. This strategy of hiding actual implementations behind a generalized QoS Marking Attribute set must be accompanied by appropriate translation and remarking functions in the originating AS border routers.

IX. RELATED WORK

A number of QoS improvement approaches have been proposed before and a selection will be briefly mentioned in this section.

Most of the approaches perform parameter signaling. [11] defines the QOS_NLRI attribute, which is used for propagating QoS-related information associated to the NLRI information conveyed in a BGP UPDATE message. Single so called "QoS routes" are signaled, which fulfill certain QoS requirements. Several information types are defined for the attribute, which concentrate on rate and delay type parameters.

[8] is based on the specified QOS_NLRI attribute and introduces some modifications to it. The notion of AS-local and extended QoS classes are distinguished, where the second group concentrates on ID associated QoS parameter propagation between adjacent peers. The first group is of more interest for this paper since it concentrates on the "identifier propagation": such as the DSCP value for example. However, this signaling is specified for the information exchange between adjacent peers only and assumes the existence of extended QoS classes and offline traffic engineering functions.

Another approach is described in [6]. It associates a list of QoS metrics with each prefix by extending the existing BGP AS_PATH attribute format. Hop-by-hop metric accumulation is performed as the AS_PATH gets extended in relaying ASes. Metrics are generically specified as a list of type length value (TLV) style attribute elements. The metrics such as bandwidth and delay are exemplary mentioned in the draft.

One contribution specialized in the signaling of Type Of Service (TOS) values which are in turn directly mapped to DSCP values in section 3.2 of the draft [25]. The TOS value is signaled within an Extended Community Attribute and, if it is understood correctly, will be applied to a certain route. An additional value field is used to identify, which routes belong to which signaled TOS community. Who advertises such attributes and whether they are of transitive or non-transitive type remains unspecified.

A comprehensive analysis is given in [2]. This "Inter-provider Quality of Service" white paper examines the inter-domain QoS requirements and derives a comprehensive approach for the introduction of at least one QoS class with guaranteed delay parameters. The implementation aspects of metering, monitoring, parameter feedback and impairment allocations are all considered in the white paper. However, QoS guarantees and parameter signaling is beyond the intention of this paper.

A very extensive work has been published in [20]. It goes far beyond this limited QoS approach of this paper. The so called “loose guarantees solution” in that work is one offered option that also renounces end-to-end QoS guarantees. However, it still performs mutual negotiations on performance parameters and bandwidth requirements. Transitive cross-layer mappings are not signaled in this approach.

Other documents may also be considered as related work as long as they convey QoS marking information, that might be "misused" for QoS class signaling.

One example is the usage of the "Traffic Engineering Attribute" as defined in IETF draft [22]. However, the attribute is non-transitive and the LSP encoding types are not generally applicable to inter-domain peering types. Its usage of the targeted QoS Marking signaling is not possible.

The second example is the current "Dissemination of flow specification rules" draft [18]. It defines a new BGP NLRI encoding format, which can be used to distribute traffic flow specifications. Such flow specification can also include DSCP values as type 11 in the NLRI. Furthermore, one could signal configuration actions together with the DSCP encoding, which could be used for filtering purposes or even trigger remarking and route selection with it. Such usage is not defined in the draft and can hardly be achieved because of the following reasons. The flow specification is focused on single flows, which might even be part of an aggregate. Such fine grained specification is counterproductive for this coarse grained general QoS marking approach. The novel approach of transitive cross-layer QoS marking could also not be incorporated, which might be essential for future tunneled
inter-domain peering.

The proposed approach of this “A Concept of inter-AS Priority Signaling using BGP Attributes” paper is now work in progress at the IETF [17].

X. BUSINESS IMPLICATIONS

Making AS-internally available traffic separations known globally, can easily lead to a traffic overload in a certain forwarding class. In order to ensure the proportional usage of the available options, this approach needs to be aligned with a locally applied class-based accounting. That is, inter-AS traffic will be counted and priced according to the available QoS class set. However, no complex service level agreements on QoS parameter boundaries and contractual penalties need to be setup for this limited QoS level of traffic separation.

REFERENCES

[1] 3GPP, “Universal Mobile Telecommunications System (UMTS); Quality of Service (QoS) concept and architecture”, 3GPP TS 23.107 version 6.2.0 Release 6, December 2004