Class of Service (CoS) in a global NGN

Zukunft der Netze – Chemnitz 2009
8. Fachtagung des ITG-FA 5.2

Thomas Martin Knoll
Chemnitz University of Technology
Communication Networks
Phone 0371 531 33246
Email knoll@etit.tu-chemnitz.de
Outline

1. Motivation
2. State of the Art
3. Proposed Improvements / Focus
4. Addressed Issues
5. Definition of new BGP Attributes
6. Implementation / Practical usage
7. Summary
## Motivation

*“Quality of Service (QoS) – what for?”*

- A generalized packet data network carries a mixture of throughput, delay and loss critical packetized information, organized in datagrams and flows.
- Approach: “Give each carried service the resources that are just right to achieve satisfactory service quality.”

---

<table>
<thead>
<tr>
<th>Motivation</th>
<th>State</th>
<th>Focus</th>
<th>Ad. Issues</th>
<th>Attributes</th>
<th>Practical Use</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Quality of Service (QoS) – what for?”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✓ A generalized packet data network carries a mixture of throughput, delay and loss critical packetized information, organized in datagrams and flows.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✓ Approach: “Give each carried service the resources that are just right to achieve satisfactory service quality.”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
State of the Art

Quality of Service (QoS) – available within domains

- Differentiated Services (DiffServ) – widely accepted and consistently enabled within administrative domains (corporate networks, ISPs, etc.)
- Traffic separation in e.g. VLANs and MPLS tunnels
- Resource reservation for tunnels – common in larger networks
  single flow reservations only on sparse (radio) links

- Marking with trust relationships
- Classification + queueing + scheduling
- Partially with reservation + admission control
State of the Art

“Quality of Service (QoS) – what for?” – the inter-domain case

- Massive over-provisioning to avoid any QoS problems (& discussions)
- Big enough to boldly request uncongested transfer links
- Uncoordinated (separate and competitive) network operation
- Trade-off: Transfer cost <-> QoS capable devices, setups, accounting, staff training, debugging, error predictability

“Throw bandwidth at the problem”
“I don’t think about QoS, I’ve got v6”
State of the Art

"Quality of Service (QoS) – what for?" – the inter-domain case

- Autonomous System -> single administrative routing domain
- AS interconnection -> peering (free) / transit (paid)
- Exchange of reachability information (routes + attributes) through BGP
- QoS within AS commonly used (IPv4/IPv6 (layer 3) combined with lower layer QoS technologies (Ethernet 802.1D, MPLS Traffic Class – RFC5462)
- NOT available (commonly used): Inter-AS QoS + Virtual Channels in lower layers for QoS → multi-parameter ingress classification / degraded quality
- Complex approaches exist, which aim for guaranteed (parameterized) QoS support for inter-AS peerings
## Proposed Improvements / Focus

### Proposed Improvements → Inter-AS CoS

- **Provides knowledge** about the available traffic separations and markings.
  - Cross-layer mapping & **transitive Cross-domain signalling** is a novel feature.
- Enables **marking adoption** (and possibly route selection) **without guarantees**.
- Fair **signalling of class overload** limitations and **excess traffic handling** with local scope.
- Greatly improves inter-AS packet forwarding.
- **Twofold “free to join” concept** (single or combined usage):
  1. global class set + cross-layer marking signalling (transitive attribute)
  2. local class set + rate limitation signalling (non-transitive attributes)

### Traffic Separation is key:

**QoS in this approach refers to primitive traffic separation into several classes, which will experience differently prioritized forwarding behaviour in relaying nodes. Enqueueing in separate queues is thereby aspired.**
**Proposed Improvements / Focus**

*Focus of the new Concept*

Distinction between AS level CoS:

1. CoS based Forwarding → use case
2. CoS based Routing → possibly future use case
3. CoS based Tunnelling → use case
**Addressed Issues**

**Cross-Layer QoS mapping**

- cross-domain tunnelling of customer traffic
  → consistent inter-layer QoS coupling
  → transparent transport

"E-LSPs"

**MPLS Label Stack Format**

- Label
- Exp.
- S
- TTL

**DS-Field**

- 0 1 2 3 4 5 6 7

**DSCP**

- Class Selector Codepoints

- Differentiated Services Codepoint
  - RFC 2474

**RFC 2474**

- ECN

**TPI**
- Tag Protocol Identifier

**VLAN TCI**
- VLAN Tag Control Information

**DATA**
- 46 .. 1500

**PAD**
- 2

**FCS**
- 4

The aim is consistent classification and a consistent class-based forwarding behaviour on all layers of an end-to-end traffic path.
Addressed Issues (cont.)

Cross-Domain QoS signalling

- Diverse usage and internal QoS strategies are not visible outside an AS
- Individual agreements (SLA) on class support between neighbouring ASes
- External BGP (eBGP) is used for Inter-Domain signalling carrying transitive CoS signalling

The aim is consistent classification and a consistent class-based forwarding behaviour on all layers of a transit traffic path.
Addressed Issues (cont.)

**CoS – Class Overload prevention**

- Traffic separation and enqueueing into separate – prioritised – queues tempts users to **overload the higher priority classes**.
- **Limitation and punishment concept** using Token Bucket filter at ingress nodes

**Motivation**

**State**

**Focus**

**Ad. Issues**

**Attributes**

**Practical Use**

**Summary**

---

*Ingress Filters to drop or down-mark excess traffic*
**BGP update message**

- **Marker**
  - 16 Bytes filled with „0xFF“

**Message Header**
- 19 octet

**Path Attributes**
- **wellknown**
  - Origin
  - AS-Path
  - Next Hop
- **optional**
  - Local Preference
  - Atomic Aggregate
  - Aggregator
  - MED
  - Community
  - Ext. Comm.

**Network Layer Routing Information**
- var. length
Definition of new BGP Attributes

Cross-domain CoS Signalling & Cross-layer Mapping


“QoS Marking Attribute” - BGP Extended Community Attribute [RFC4360]

Signalling of ‘original & active’ traffic class markings for several layers across ASes using transitive extended communities with IANA assigned type 0x04.

Motivation    State    Focus    Ad. Issues    Attributes Practical Use    Summary
Definition of new BGP Attributes

CoS – Class Overload prevention


“CoS Capability Attribute” - BGP Extended Community Attribute [RFC4360]

Signalling of supported traffic classes → currently limited to LE, BE, AF, EF using non-transitive extended communities with IANA assigned type 0x40.
Definition of new BGP Attributes

CoS – Class Overload prevention

<table>
<thead>
<tr>
<th>Bit</th>
<th>Flag</th>
<th>Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>BE</td>
<td>Default to ‘1’ to signal general “Best Effort” PHB support</td>
</tr>
<tr>
<td>1</td>
<td>EF</td>
<td>‘1’ … “ Expedited Forwarding” PHB support [10]</td>
</tr>
<tr>
<td>2</td>
<td>AF</td>
<td>‘1’ … “ Assured Forwarding” PHB group support [11]</td>
</tr>
<tr>
<td>3</td>
<td>LE</td>
<td>‘1’ … “ Lower Effort” PHB support [5]</td>
</tr>
<tr>
<td>4..7</td>
<td>unused</td>
<td>Default to ‘0’</td>
</tr>
</tbody>
</table>

Class Set selection
→ 2 classes – BE + LE  
→ 3 classes – BE + EF + AF  
→ 4 classes – BE + LE + EF + AF

Motivation | State | Focus | Ad. Issues | Attributes | Practical Use | Summary
---|---|---|---|---|---|---

Definition of new BGP Attributes

CoS – Class Overload prevention

The new CoS Parameter Attribute is a variable length non-transitive attribute, which is not readily available as yet.

G flag … globally or NLRI local

DR flag… drop / remarking
Implementation / Practical usage

- Linux BGP routing (Quagga) has been augmented with the selective Cross-domain & Cross-Layer CoS Marking scheme.

- The Network Analyzer “Wireshark” has been officially extended to detect and interpret the new BGP extended community attributes.

- Lab tests with Cisco routers have been performed, which enabled feasibility testing, statistical calculations on real world resource usage and revealed the output of the new attribute information within Debug logs.

- An online decoding service for Cisco debug logs is available at: http://www.bgp-qos.org/draft-knoll/decode_attributes.php
Tests with Internet Exchange Points have been performed in order to document their CoS support (IEEE 802.1p) on the switching platform.
Implementation / Practical usage


Details of Internet Exchange Point: **DE-CIX**

<table>
<thead>
<tr>
<th>Name</th>
<th>DE-CIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL</td>
<td><a href="http://www.de-cix.net/">http://www.de-cix.net/</a></td>
</tr>
<tr>
<td>Country</td>
<td>DE</td>
</tr>
<tr>
<td>First Name</td>
<td>Arnold</td>
</tr>
<tr>
<td>Surname</td>
<td>Nipper</td>
</tr>
<tr>
<td>E-mail</td>
<td><a href="mailto:arnold.nipper@de-cix.net">arnold.nipper@de-cix.net</a></td>
</tr>
<tr>
<td>Phone</td>
<td>+49 69 1730 902 -0</td>
</tr>
<tr>
<td>Fax</td>
<td>+49 69 4056 2716</td>
</tr>
</tbody>
</table>

Details of Internet Exchange Point: **AMS-IX**

<table>
<thead>
<tr>
<th>Name</th>
<th>AMS-IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>NL</td>
</tr>
<tr>
<td>First Name</td>
<td>AMS-IX NOC</td>
</tr>
<tr>
<td>Surname</td>
<td>AMS-IX NOC</td>
</tr>
<tr>
<td>E-mail</td>
<td><a href="mailto:noc@ams-ix.net">noc@ams-ix.net</a></td>
</tr>
<tr>
<td>Phone</td>
<td>+31 (20) 305 89 99</td>
</tr>
<tr>
<td>Fax</td>
<td>+31 (20) 305 89 90</td>
</tr>
</tbody>
</table>

**DE-CIX** supports the following attributes:

- IEEE 802.1p support on request + non-blocking line-rate switching
- QoS VLAN
- QoS VLAN priority

**AMS-IX** supports the following attributes:

- IEEE 802.1p support + non-blocking line-rate switching
- QoS VLAN
- QoS VLAN priority

**DE-CIX** has 8 QoS queues, while **AMS-IX** has 4 QoS queues (7 later in 2009).
Summary

- The proposed approach enables a general QoS based forwarding which allows for informed routing and marking decisions. It is optimized for ease of deployment and adopted to the current poor inter-domain forwarding model.

- The concept aims for a consistent and widely adopted QoS approximation, which encompasses cross-layer and cross-domain traffic class handling from L1 to at least L3 as generally offered QoS treatment.

- The concept incorporates a confidentiality option that allows operators the distinction between an secluded internal and the advertised external Class Set.

- More sophisticated QoS concepts are not prohibited and will always exist, which results in future “better quality islands/path”.
Thank you for your attention.