IP QoS DiffServ

Differentiated Services Architecture
DSCP, CAR

Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

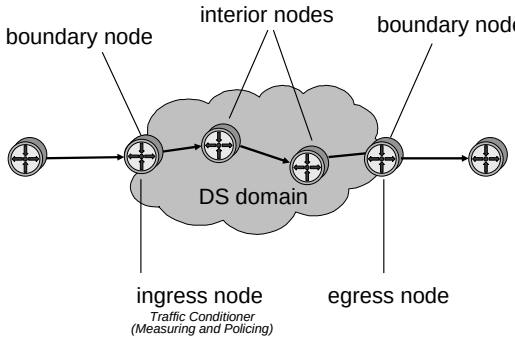
Principles

- **Integrated Services Model does not scale well**
  - flow based
  - traffic overhead (RSVP messages)
  - routers must maintain state information for each flow
- **ATM´s principle of traffic classes**
  - seemed to be also useful for IP
  - note: ATM virtual circuits implicitly identify flows (connection between two end systems) but queuing management handels service classes only (CBR, VBR, ABR, UBR) -> all virtual circuits of same class use the same queue
- **Idea of Differentiated Services Model (DS, DiffServ): Packets are separated into traffic classes**
  - aggregation of traffic → better scalability!
- **Routers treat each traffic class according specific service level agreements**

Principles

- **Requires a traffic identification and labeling (marking) mechanism**
- **Routers examine this label to adjust queuing and drop parameters**
  - IP Type of Service (ToS) field used as DS label
  - now called "Differentiated Service Code Point" (DSCP)
- **So there are two main DS building blocks:**
  - DSCP
  - Per-Hop Behavior (PHB) of the routers
- **No reservations necessary**
  - static QoS mechanism
  - PHB is determined by DSCP

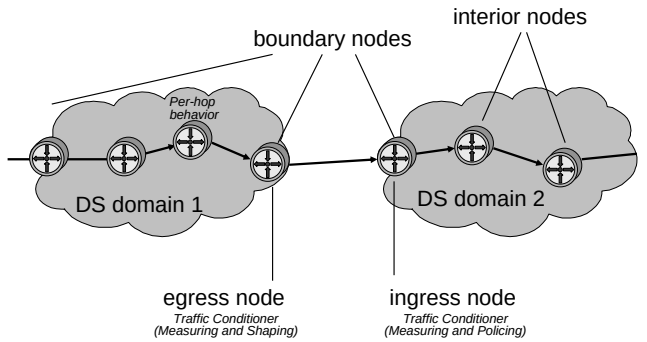
DS Terminology



DS Architecture

- **Distinction between <u>boundary</u> and <u>interior</u> functions/nodes**
- **DSCP bits must be set at network boundaries**
  - at “ingress” point to achieve PHP in own DS domain
  - at “egress” points to conform to DSCP meaning in next DS domain administered by another authority
- **Also conditioning of packets at network boundaries**
  - at “ingress” point for traffic policing
  - at “egress” points for traffic shaping to conform traffic aggrement with next DS domain administered by another authority
- **Interior nodes have to examine these bits and treat the packets accordingly**
  - interior nodes schedule packet forwarding in different ways based on DSCP

DS Multiple Domains



DS

- **DS requires routers to support sophisticated queue management**
  - DSCP value specifies queue parameters and drop preferences
  - traditional FIFO queuing cannot provide service differentiation
- **DS assumes a “Service Level Agreement” between adjacent networks**
  - service provider and customer
  - between service providers
    - adapt different per-hop behaviors

Service Level Agreement / Specification

- **Service Level Agreement (SLA)**
  - in context to DS is a contract between customer and service provider that specifies the forwarding service
  - non-technical aspects
    - pricing
    - contractual obligations
- **Service Level Specification (SLS)**
  - a set of parameters and values
  - defines the service offered to a traffic stream by a DS domain
  - Traffic Conditioning Specification (TCS)
    - specifies a set of classifier rules

Traffic Conditioning Agreement

- **A Traffic Conditioning Agreement (TCA) consists**
  - classifier rules
  - traffic profiles (optional)
    - bandwidth, throughput, latency, drop precedence,…
  - metering, marking, discarding, shaping rules
- **A TCA is the result**
  - of an SLA
  - the specific service requirements
  - and the providers service provisioning policy

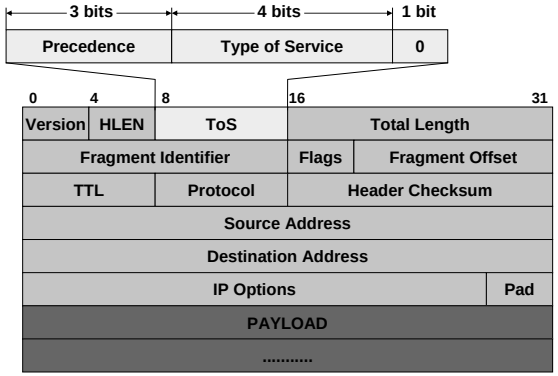
Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

Providing a Label for DS

- **In earlier IP-days the IETF intended the Type of Service (ToS) field in the IP header to be used as some kind of flow descriptor**
  - rarely used, mostly even ignored by operating systems
- **So, this field can be reused !**
  - no additional header necessary
- **The IPv4 ToS field consists of**
  - three precedence bits for eight precedence levels (0...lowest, 7...highest)
  - four type of service indication bits
  - one unused bit

IPv4 Header Review



IPv4 / IPv6 ToS Recycling

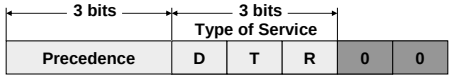
- **IPv4 ToS field was redefined by the IETF to become the "Differentiated Service Code Point" (DSCP)**
- **Now the DSCP field is used to label the traffic class of a flow**
- **In IPv6 the "Traffic Class" octet is used to implement DiffServ**

Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

IPv4 ToS: Historical Review

- **1981: RFC 791 defines the Internet Protocol containing the original ToS-Field**

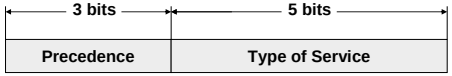


Precedence:
111 - Network Control
110 - Internetwork Control
101 - CRITIC/ECP
100 - Flash Override
011 - Flash
010 - Immediate
001 - Priority
000 - Routine

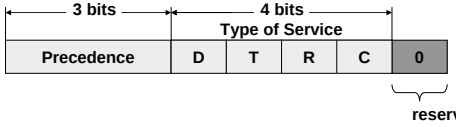
Type of Service:
D: delay (0 = normal / 1 = low)
T: throughput (0 = normal / 1 = high)
R: reliability (0 = normal / 1 = high)

IPv4 ToS: Historical Review

- **1989: RFC 1122 defines ToS as a five bit field**



- **1992: RFC 1349 redefines the ToS-Field to be four bits**

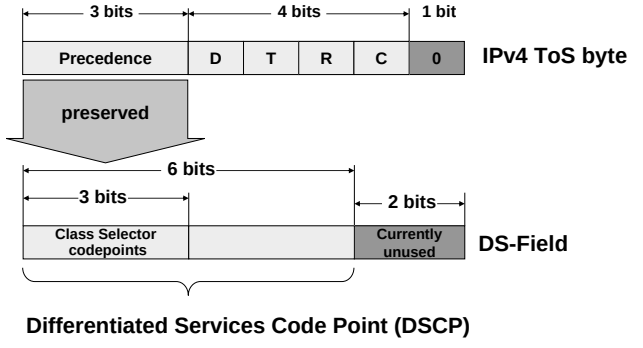


Initial Meaning of the ToS Subfield

Type of Service	Bit Flags	Value	Examples
Low Latency	1000	8	Telnet keystrokes, urgent data, etc.
High Throughput	0100	4	FTP downloads, backups, bandwidth-sensitive applications
High Reliability	0010	2	File-sharing, database updates, UDP transactions
Low Cost	0001	1	NNTP news feed, nonessential traffic
Default	0000	0	Normal traffic

IPv4 ToS → DSCP

- **1998: RFC 2474 specifies DS Field**

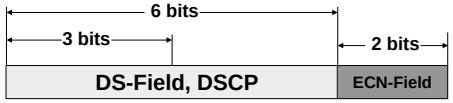


Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

IPv4 ToS → newest development: ECN

- **2002: RFC 3260 declares Explicit Congestion Notification (ECN) as an integration of the former ToS-Field**



0	0	Not-ECT
0	1	ECT(1)
1	0	ECT(0)
1	1	CE

ECT ECN-Capable Transport
CE Congestion Experienced

ECN in detail

- **ECT(0) or ECT(1) are set by the data sender to indicate a ECN-capability**
  - routers treat both code points as equivalent
- **Not-ECT indicates a packet that is not using ECN**
- **CE code point is set by a router to indicate congestion to the end nodes**
- **Routers that have a packet arriving at a full queue drop the packet, just as they do in the absence of ECN**
  - see specifications
    - RFC 3168 "The Addition of ECN to IP"

DSCP Usage

- **DS-routers overwrite the IPv4 ToS field**
  - only precedence bits are preserved
  - now known as "class selector code point" field
- **DSCP consists of 6 bits**
  - 64 code points possible:
    - 48 in global space, 16 for local use
  - last 2 bits of the former IPv4 ToS field used for ECN
  - specifications
    - RFC 2474: "Definition of the Differentiated Service Field in the IPv4 and IPv6 Headers"
    - RFC 2475: "An Architecture for Differentiated Services"
    - RFC 3260: "New Terminology and Clarifications for DiffServ"

Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

Per-Hop Behavior

- **A Per-Hop Behavior (PHB) is**
  - "The externally observable forwarding behavior applied at a DS-compliant node to a DS behavior aggregate."
  - concerns local behavior of a network-node
  - should enable predictable services
- **The PHB describes various node-behaviors for incoming packets**
  - forwarding
  - classification
  - scheduling
  - drop
- **Note: Traffic conditioning and service provisioning is not part of a PHB definition**

Per-Hop Behaviors

4 Types:

- **Default PHB (DE)**
  - Best Effort
  - DSCP = 000000
- **Class Selector PHB**
  - Defined to be backward-compatible with IP precedence

Precedence 1	**001 000**
Precedence 2	**010 000**
Precedence 3	**011 000**
Precedence 4	**100 000**
Precedence 5	**101 000**
Precedence 6	**110 000**
Precedence 7	**111 000**

Per-Hop Behaviors

- **Expedited Forwarding (EF)**
  - Premium Service (Highest DS QoS)
  - using a single DSCP only (101110)
  - minimal delay and jitter; low loss and assured bandwidth
  - targets applications like VoIP and video conferencing
    - virtual leased line service
    - looks like a point-to-point connection for DiffServ network's end nodes
  - implementation by appropriate packet scheduling techniques
    - it must be ensured, that a busy EF queue does not starve the remaining traffic queues
    - e.g. CBWFQ (Class-Based Weighted Fair Queuing)
    - e.g. Priority Queuing

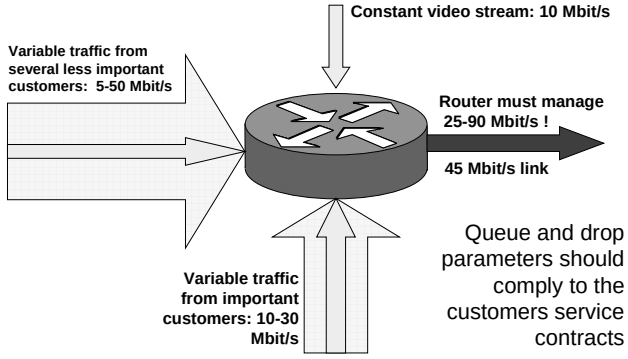
Per-Hop Behaviors

- **Assured Forwarding (AF)**
  - emulates a lightly loaded network even in congestion cases
  - 4 classes and 3 drop-precedence's within each class = 12 code points
    - bandwidth assurance but no guarantee
    - each traffic class is serviced in its own queue

Drop Precedence	Class 1	Class 2	Class 3	Class 4
Low	**001 010**	**010 010**	**011 010**	**100 010**
Medium	**001 100**	**010 100**	**011 100**	**100 100**
High	**001 110**	**010 110**	**011 110**	**100 110**

  - Random Early Detection (RED)-queue management is often used for an implementation

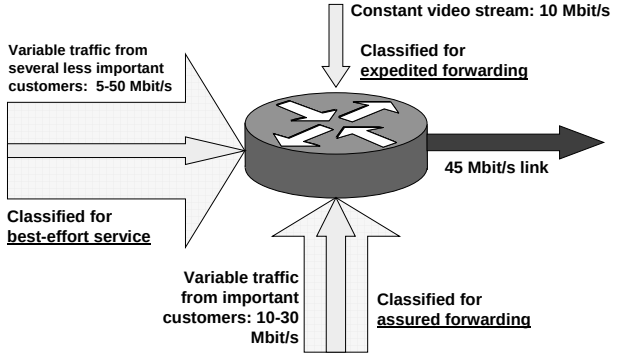
Solution: Traffic Differentiation



Queue and drop parameters should comply to the customers service contracts

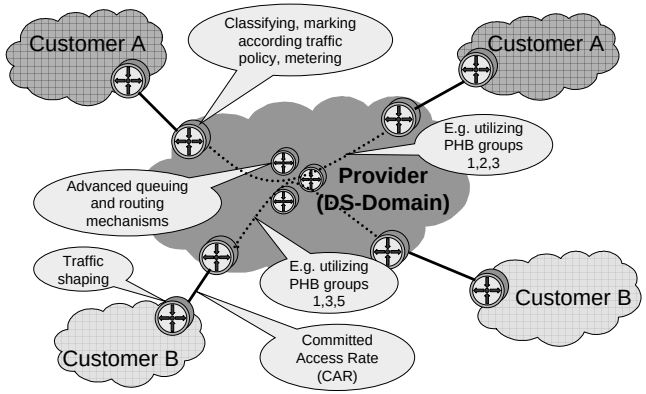
 IP QoS DiffServ, v4.3 29

Internet: Competitive Traffic



 IP QoS DiffServ, v4.3 30

DS in Action



 IP QoS DiffServ, v4.3 31

Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

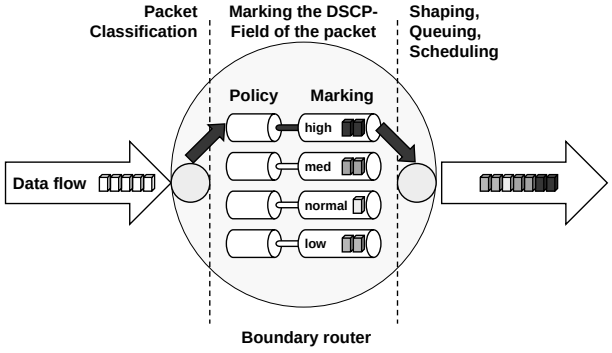
 IP QoS DiffServ, v4.3 32

DS Domains

- **A “DS domain” is a contiguous subnet with a consistent DS management**
- **A "DS region" consists of one or more contiguous DS domains**
- **Related PHB´s are collected in a “PHB-group”**
  - because there may be many more potential PHB´s than the 64 map-able PHB´s that can be addressed by the DSCP
  - e.g. several PHB´s with similar properties can be summarized as one specific PHB group
    - sharing similar properties such as buffer sizes, bandwidth, delay, jitter, loss, etc.

Boundary nodes

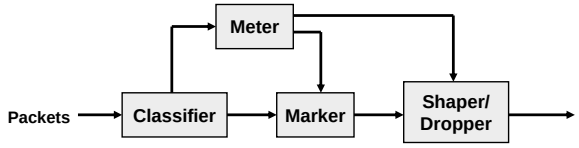
- **Traffic Conditioning**



Traffic Profiles

- **Classifiers check if incoming traffic profile complies to the SLA´s traffic profile**
- **A traffic profile can be given by token-bucket parameters**
  - eg. "DSCP=x, use token-bucket r, b" specifies that all packets with DSCP=x should be measured against a token bucket meter with rate r and burst size b.
  - out-of-profile traffic: when packets arrive while there is currently no token available
- **Solution against out-of-profile traffic:**
  - DSCP remarking
  - traffic shaping: queuing to adapt token rate
  - packet dropping

Traffic Conditioner



Traffic conditioners are usually located within DS ingress and egress boundary nodes

Traffic Conditioner

- **Classifier**
  - selects a packet in a traffic stream based on the content of some part of the packet header
    - common way: classification based on the DSCP field
- **Meter**
  - measures traffic stream against traffic profile
    - based on traffic descriptor such as a token bucket
  - the state of the Meter influences the decision of the Marker and the Shaper/Dropper

Traffic Conditioner

- **Marker**
  - might be configured to mark all packets which are steered to it (by the classifier) to a single DSCP
  - might select a PHB in a PHB group according to the meter´s state
  - might re-mark (change an existing DSCP)
- **Shaper**
  - delays some or all packets in a traffic stream to smooth bursts and cares for compliance with the specified traffic profile
- **Dropper**
  - discards all packets that doesn't comply with the traffic profile ("policing the stream")

Packet Classification

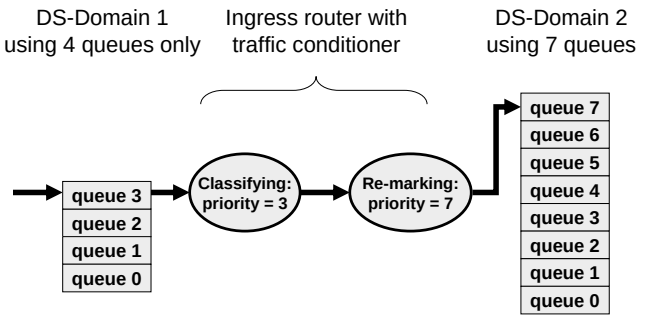
Two types of classifiers:

- **Behavior Aggregate (BA) classifier**
  - classifies packets based on the DSCP only
- **Multi-field (MF) classifier**
  - selects packets based on the value of a combination of one or more header fields
  - e.g. source and destination address, DSCP, protocol ID, ports, also incoming interface
  - note: IP fragmentation can lead to wrong MF classification of fragments if TCP fields are examined

Re-marking the DSCP

- **Ingress routers must ensure that packets entering a DS domain receive the same QoS as in the domain before**
  - extending the PHB group by re-marking the DSCP field and traffic conditioning
- **Ingress router examines the TCP port numbers and assigns a new service class**
- **So, the DS is transparent to applications**
  - note: IntServ requires application's help !
- **However, an end-to-end QoS cannot be guaranteed**
  - because no reservation of resources
  - "Soft" QoS

DSCP Re-marking



Source Domain

- **The source domain is the place where the traffic originates**
  - Usually the customer´s domain
- **Nodes within a source domain may perform**
  - traffic classification
    - E.g. directly by the source node
    - "initial marking" or "pre-marking"
  - conditioning
- **Advantages of pre-marking and early conditioning:**
  - traffic source knows preferences of applications
  - e.g. a CEO´s traffic has more importance
  - simplifies classification and traffic shaping tasks on border nodes
- **However, source domain is responsible that aggregated traffic conforms to the TCA**
  - signaling of PHB´s using RSVP might be necessary

Agenda

- **DiffServ Principles**
- **DS-Field, DSCP**
  - Historical Review
  - Newest Implementations
- **Per-Hop Behaviors (PHB)**
- **DiffServ in Detail**
- **DiffServ in other Environments**

Multicast Traffic

- **IP multicast traffic may cause a higher load inside a DS domain**
  - multicast packets may take multiple paths simultaneously because of multicast packet replication
  - unpredictably consumes more network resources than unicast packets
- **So it may be necessary to reserve different sets of DSCP for unicast and multicast traffic**
  - to provide resource isolation from multicast traffic

QoS Mapping between IEEE 802 and IP

- **802.1p and 802.1Q define 8 levels of prioritization (0-7)**
- **DSCP class selector code point also provides 8 precedence levels**
  - three-bit field within DSCP
- **So 802.1p priority values can be mapped one-to-one into the DS class selector code point**

DS and IPsec

- **IPsec does not use the DS field of the IP header for its cryptographic calculations**
  - That is: IPsec can be used in DS networks without modifications
- **DS together with IPsec tunneling hides port information**
  - increases security
  - as routers only need to examine the DSCP to perform traffic differentiation

DS and IPsec

IPsec tunnel mode:

- **encapsulates the whole IP packet with an outer IP header**
- **outer IP header contains a proxy-DS field which is set according to the SLA**
- **only this outer proxy-DS field will be modified on the way through the IPsec tunnel**
- **at the end of the IPsec tunnel the outer IP header is stripped off again**