IP Differentiated Services

- Providing differentiated services in IP networks can be a challenge because of the scaling problems presented by the large number of flows present in large networks makes the cost of maintaining per-flow state unacceptable
- Moreover, operators need the ability to perform *traffic engineering* by directing flows of traffic along specific paths
- The IP Differentiated Services (DiffServ) architecture is a way to obtain a scalable IP network with QoS guarantees

IP Differentiated Services, basic concepts

- The IP DiffServ architecture defines a framework within which operators/ISPs may provide end-to-end differentiated services in a coordinated and reliable fashion
- With such an architecture, an operator/ISP would be able to craft common agreements for the handling of differentiated services in a consistent fashion, facilitating end-to-end differentiated services via a composition of these agreements
- Supporting differentiated services is a strategic objective for an operator, as support would allow it to offer special services such as priority for bandwidth for mission critical services for users willing to pay a service premium
- Customers would contract with ISPs for these services under Service Level Agreements (SLAs)
- Such an agreement may specify the traffic volume, how the traffic is handled, and the associated billing policy

IP Differentiated Services, basic concepts

- Another important component in the DiffServ architecture is the advent of *traffic engineering*
- Traffic engineering is the ability to move trunks away from the path selected by the ISP's IGP (Interior gateway protocol, such as RIP, OSPF, EIGRP, ...) and onto a different path
- This allows an ISP to route traffic around known points of congestion in its network, thereby making more efficient use of the available bandwidth
- In turn, this makes the ISP more competitive within its market by allowing the ISP to pass lower costs and better service on to its customers

Traffic classes

- Traffic flows may fall into a variety of different traffic classes
- For ISP operations, it is essential that packets be accurately classified before entering the ISP and that it is very easy for an ISP device to determine the traffic class for a particular packet
- The traffic class of MPLS packets can be encoded in the CoS field within the MPLS label header
- In addition, traffic classes for IPv4 packets can be classified via the IPv4 ToS byte, called DSCP (*Differentiated Services Code Point*) in the DiffServ framework
- Classification of packets at the network ingress allows for a differentiated treatment of packets inside the network

Main features of the IP DiffServ architecture

- The main purpose of the IP DiffServ architecture is to implement scalable service differentiation
- This architecture achieves scalability by aggregating traffic classification state which is conveyed by means of IP-layer packet marking using the DS field [DSFIELD]
- Packets are classified and marked to receive a particular per-hop forwarding behavior on nodes along their path
- Sophisticated classification, marking, policing, and shaping operations need only be implemented at network boundaries or hosts
- Network resources are allocated to traffic streams by service provisioning policies which govern how traffic is marked and conditioned upon entry to a differentiated services-capable network, and how that traffic is forwarded within that network

Main features of the IP DiffServ architecture

- The architecture is composed of a number of functional elements implemented in network nodes, including
  - a small set of per-hop forwarding behaviors
  - packet classification functions
  - Traffic conditioning functions including
    - Metering
    - Marking
    - Shaping
    - Policing
- This architecture achieves scalability by implementing complex classification and conditioning functions only at network boundary nodes, and by applying per-hop behaviors to aggregates of traffic which have been appropriately marked using the DS field in the IPv4 or IPv6 headers [DSFIELD]
- Per-hop behaviors are defined to permit a reasonably granular means of allocating buffer and bandwidth resources at each node among competing traffic streams

Main features of the IP DiffServ architecture

- Per-application flow or per-customer forwarding state need not be maintained within the core of the network
- A distinction is maintained between:
  - the service provided to a traffic aggregate
  - the conditioning functions and per-hop behaviors used to realize services
  - the DS field value (DS codepoint) used to mark packets to select a per-hop behavior
  - the particular node implementation mechanisms which realize a per-hop behavior

Differentiated Services domain

- A DS domain is a contiguous set of DS nodes which operate with a common service provisioning policy and set of PHB groups implemented on each node
- A DS domain has a well-defined boundary consisting of DS boundary nodes which classify and possibly condition ingress traffic to ensure that packets which transit the domain are appropriately marked to select a PHB from one of the PHB groups supported within the domain
- Nodes within the DS domain select the forwarding behavior for packets based on their DS codepoint, mapping that value to one of the supported PHBs using either the recommended codepoint->PHB mapping or a locally customized mapping [DSFIELD]

Differentiated Services domain

- A DS domain normally consists of one or more networks under the same administration
- For example, an organization's intranet or an ISP
- The administration of the domain is responsible for ensuring that adequate resources are provisioned and/or reserved to support the SLAs offered by the domain
- Both DS boundary nodes and interior nodes must be able to apply the appropriate PHB to packets based on the DS codepoint
- Interior nodes may be able to perform limited traffic conditioning functions such as DS codepoint re-marking
- A host in a network containing a DS domain may act as a DS boundary node for traffic from applications running on that host
- If a host does not act as a boundary node, then the DS node topologically closest to that host acts as the DS boundary node for that host's traffic

DS boundary nodes

- DS boundary nodes act both as a DS ingress node and as a DS egress node for different directions of traffic
- Traffic enters a DS domain at a DS ingress node and leaves a DS domain at a DS egress node
- A DS ingress node is responsible for ensuring that the traffic entering the DS domain conforms to its TCA
- A DS egress node may perform traffic conditioning functions on traffic forwarded to a directly

Differentiated Services Region

- A differentiated services region (DS Region) is a set of one or more contiguous DS domains
- DS regions are capable of supporting differentiated services along paths which span the domains within the region
- The DS domains in a DS region may support different PHB groups internally and different codepoint->PHB mappings
- However, to permit services which span across the domains, the peering DS domains must each establish a peering SLA which a TCA which specifies how transit traffic from one DS domain to another is conditioned at the boundary between the two DS domains
- It is possible that several DS domains within a DS region may adopt a common service provisioning policy and may support a common set of PHB groups and codepoint mappings, thus eliminating the need for traffic conditioning between those DS domains

Traffic Classification and Conditioning across domains

- Differentiated services are extended across a DS domain boundary by establishing a
- The SLA may specify packet classification and re-marking rules and may also specify traffic profiles and actions to traffic streams which are in- or out-of-profile
- The TCA between the domains is derived (explicitly or implicitly) from this SLA
- The packet classification policy identifies the subset of traffic which may receive a differentiated service by being conditioned and/or mapped to one or more behavior aggregates
- Traffic conditioning performs metering, shaping, policing and/or re-marking to ensure that the traffic entering the DS domain conforms to the rules specified in the TCA

Classification

- Packet classifiers select packets in a traffic stream based on the content of some portion of the packet header
- There are two types of classifiers
  - The BA (Behavior Aggregate) Classifier classifies packets based on the DS codepoint only
  - The MF (Multi-Field) classifier selects packets based on the value of a combination of one or more header fields, such as source address, destination address, DS field, protocol ID, source port and destination port numbers, and other information
- Classified packets can be subsequently conditioned

Traffic profile

- A traffic profile specifies the statistical properties of a classified traffic stream
- It provides rules for determining whether a particular packet is in-profile or out-of-profile
- For example, a profile based on a token bucket may look like:
  - if codepoint=X, then use token-bucket r, b
- The above profile indicates that all packets marked with DS codepoint X should be measured against a token bucket meter with rate *r* and burst size *b*

Traffic profile

- Different conditioning actions may be applied to the in-profile packets and out-of-profile packets
- In-profile packets may be allowed to enter the DS domain without further conditioning; or, alternatively, their DS codepoint may be changed
- The latter happens, for example, when the packets enter a DS domain that uses a different PHB group or codepoint->PHB mapping policy for this traffic stream
- Out-of-profile packets may be queued until they are in-profile (shaped), discarded (policed), marked with a new codepoint (re-marked, mapped to one behavior aggregates that is "inferior" in some dimension of forwarding performance to the BA into which in-profile packets are mapped

Per Hop Behaviors

- A per-hop behavior (PHB) is a description of the forwarding behavior of a DS node applied to a particular DS behavior aggregate
- Multiple behavior aggregates may compete for buffer and bandwidth resources on a node
- The PHB is the means by which a node allocates resources to behavior aggregates, and it is on top of this basic hop-by-hop resource allocation mechanism that useful differentiated services may be constructed
- Schedulers play a key role in the practical implementation of a PHB

Per Hop Behaviors

- For example, a PHB may guarantee a minimal bandwidth allocation of X% of a link
- The implementation of such PHB naturally calls for a rate-based scheduler
- Alternatively, a PHB would guarantee a minimal bandwidth allocation of X% of a link, with proportional fair sharing of any excess link capacity (see the GPS scheduler)
- PHBs may be specified in terms of their priority relative to other PHBs (see the static priority scheduler), or in terms of their relative observable traffic characteristics (e.g., delay, see he Earliest Deadline First scheduler)

Per Hop Behaviors

- PHBs may be used as building blocks to allocate resources and should be specified as a group (PHB group) for consistency
- PHB groups share a common constraint applying to each PHB within the group, such as a packet scheduling or buffer management policy
- The relationship between PHBs in a group may be in terms of absolute or relative priority (e.g., service priorities and/or discard priorities)

Per Hop Behaviors

- PHBs are implemented in nodes by means of some buffer management and packet scheduling mechanisms
- PHBs are defined in terms of behavior and not in terms of the practical implementation
- In fact, a variety of implementation mechanisms may be suitable for implementing a particular PHB group and this is under the responsibility of the provider
- A PHB is selected at a node by a mapping of the packet's DS codepoint
- Standardized PHBs have a recommended codepoint. However, the total space of codepoints is larger than the space available for recommended codepoints for standardized PHBs, and [DSFIELD] leaves provisions for locally configurable mappings

SLAs and TCAs in the IP DiffServ architecture

- The Diffserv Architecture uses the term *Service Level Agreement* (SLA) to describe the "service contract... that specifies the forwarding service a customer should receive“
- The SLA may include traffic conditioning rules which constitute a *Traffic Conditioning Agreement* (TCA)
- A TCA is "an agreement specifying classifier rules and any corresponding traffic profiles and metering, marking, discarding and/or shaping rules which are to apply....“
- The notion of an "agreement" implies also considerations about pricing, contractual issues and other business issues
- There also could be other technical considerations in such an agreement (e.g., service availability) which are not addressed by Diffserv

SLAs and TCAs in the IP DiffServ architecture

- The notions of SLAs and TCAs should be taken to represent the broader context, beyond strictly technical aspects
- A *Service Level Specification* (SLS) is a set of parameters and their values which together define the service offered to a traffic stream by a DS domain
- A *Traffic Conditioning Specification* (TCS) is a set of parameters and their values which together specify a set of classifier rules and a traffic profile
- A TCS is an integral element of an SLS
- A *traffic stream* can be an individual microflow or a group of microflows or it can be a BA
- Thus, an SLS may apply in the source or destination DS domain to a single microflow or group of microflows, as well as to a BA in any DS domain

SLAs and TCAs in the IP DiffServ architecture

- A *Service Provisioning Policy* is "a policy which defines how traffic conditioners are configured on DS boundary nodes and how traffic streams are mapped to DS behavior aggregates to achieve a range of services."
- According to RFC 3198, a policy is "...a set of rules to administer, manage, and control access to network resources"
- Therefore, the relationship between an SLS and a service provisioning policy is that the latter is, in part, the set of rules that express the parameters and range of values that may be in the former

PHB groups

- A *PHB group* is "a set of one or more PHBs that can only be meaningfully specified and implemented simultaneously, due to a common constraint applying to all PHBs in the set such as a queue servicing or queue management policy
- A PHB group provides a service building block that allows a set of related forwarding behaviors to be specified together (e.g., four dropping priorities)
- A single PHB is a special case of a PHB group

Differentiated Services Field Definition

- A replacement header field, called the DS field, is defined, which is intended to supersede the existing definitions of the IPv4 TOS octet [RFC791] and the IPv6 Traffic Class octet [IPv6]
- Six bits of the DS field are used as a codepoint (DSCP) to select the PHB a packet experiences at each node
- A two-bit currently unused (CU) field is reserved
- The value of the CU bits are ignored by differentiated services-compliant nodes when determining the per-hop behavior to apply to a received packet