Helsinki University of Technology 17.10.2001
Mat-2.108 Independent research projects in applied mathematics

# Analytical Model of AF PHB Node in DiffServ Network

**Eemeli Kuumola**
47021T
TFY
e-mail: ekuumola@cc.hut.fi

# Analytical Model of AF PHB Node in DiffServ Network

## Contents

**Abstract**

The current Internet offers only best-effort service for its users. Increase in the number of Internet users and the introduction of new Internet based services has increased the need for a Internet service quality differentiation. DiffServ architecture is one of the suggested methods to provide QoS into the Internet. In this report, a model is constructed for an Assured Forwarding (AF) PHB node in a DiffServ network with TCP and UDP traffic. The model describes two AF classes with three drop precedence levels. The parameters for AF packet classifier mechanism are adapted from the Simple Integrated Media Access (SIMA) nominal bit rate mechanism. The model is used in studying how the differentiation in bandwidth allocation between flows can be achieved with the AF PHB mechanism. The numerical results show that with AF PHB mechanism differentiation can be achieved. The bandwidth allocation ratio is not fixed, but depends on the network load. The results also show that by marking the TCP and UDP flows into separate AF classes, congestion sensitive TCP flows can be protected from bandwidth exhaustion of the UDP flows.

# 1 Introduction

The current Internet is based on a technology that provides best-effort service without specific performance guarantees for the individual end user. All the data transmitted over the Internet is treated in the same way regardless of the importance aspects the user would like to set for the data. The situation could be compared to a postal service offering only one type of delivery without an option to choose between different urgency or delivery assurance classes. So far the best-effort service has been adequate for the most of the Internet users. However, both the steadily growing number of the Internet users and the introduction of new Internet based applications have increased a need for a Internet service differentiation. A workable method for the service differentiation would also lay a base for differentiated pricing of Internet services.

Originally the scope of the Internet Protocol was to define functions necessary to deliver data packets from a source to a destination over an interconnected system of networks without providing any kind of end-to-end data transmission reliability. Although the Internet Protocol supports a quality of service marking in the IP packet, it does not define any policy for service differentiation [11]. Recently research has been carried out to create a simple scheme that would provide a range of service levels in the Internet. The Internet Engineering Task Force (IETF) has suggested several approaches for the implementation of Quality of Service (QoS) in the Internet, though none of the approaches have yet been approved for the basis of the future Internet. The most notable proposals for the Internet QoS mechanism are the Integrated Service/Resource Reservation Protocol method (IntServ/RSVP) suggested in [12] and the Differentiated Services (DiffServ) architecture presented in [1]. The IntServ/RSVP introduced a service differentiation mechanism that is based on flow control and flow level scheduling inside the network. Due to excessive computational and memory requirements that are needed to control a large amount of data flows at the high speed links, the IntServ/RSVP method is poorly scalable to larger networks [13]. In the DiffServ method the scalability problem is bypassed by setting the sophisticated flow classification operations in the edges of the network and implementing the QoS inside the network with simple packet level Per-Hop Behavior (PHB) mechanisms. So far the DiffServ architecture is considered to be a more promising alternative for the implementation of QoS in the Internet.

An issue that still needs more research in the DiffServ architecture is how the flow level quality requirements and fairness issues, e.g., packet loss probabilities and fairness in bandwidth allocation, can be achieved by means of the packet level control mechanisms. Functioning of various DiffServ PHB groups, including Assured Forwarding (AF) and Expedited Forwarding (EF), has been analyzed with analytical models in, e.g., [9] and with simulation models in, e.g., [3]. More profound analysis on relation of flow level requirements in a DiffServ network with Simple Integrated Media Access (SIMA) PHB is presented in [10].

In this report functionality of an AF PHB node in a DiffServ network with TCP and UDP traffic is studied with an analytical model. The model is used in analyzing how differentiated bandwidth allocation between traffic flows is achieved with the AF PHB mechanism, and how the congestion sensitive TCP flows can be protected from the bandwidth exhaustion of the UDP flows.

The structure of this report is following. First the DiffServ architecture and Assured Forwarding PHB group is introduced briefly. After this, an analytical model for the DiffServ network with an AF PHB node and TCP and UDP traffic is presented. Finally, the results gained from the model are shown. Based on the results, the AF PHB mechanism can achieve differentiation in bandwidth allocation between the flows. However, the bandwidth allocation ratio is not fixed, but depends on the network load. Also, when the TCP and UDP flows are separated into different AF classes, TCP flows can be protected from the UDP flows.

# 2 Differentiated Service Architecture

The Differentiated Service architecture proposed in [1] defines a framework for implementing scalable service differentiation in the Internet. The DiffServ architecture is composed of several mechanisms implemented in the network nodes. These mechanisms include traffic conditioning functions, packet classification functions and a set of per-hop forwarding behaviors. The scalability of the DiffServ mechanism is achieved by implementing complex classification and conditioning functions only at the boundary network nodes and applying simple per-hop behaviors in the interior network nodes.

The top level DiffServ architecture is illustrated in Figure 1. The network is constructed of end-user terminals that are connected to boundary nodes and further to core nodes that forward traffic in the interior parts of a DiffServ network. At the boundary nodes temporal properties of traffic flow are measured and compared against a predefined traffic profile. Packets are classified into few Per-Hop Behavior classes based on this flow level comparison. Before sending the packets into the interior parts of network, the packets may be delayed or dropped to shape traffic flows to bring them in compliance with corresponding traffic profiles. Packets that are marked into the same PHB class will experience similar forwarding behavior in the core

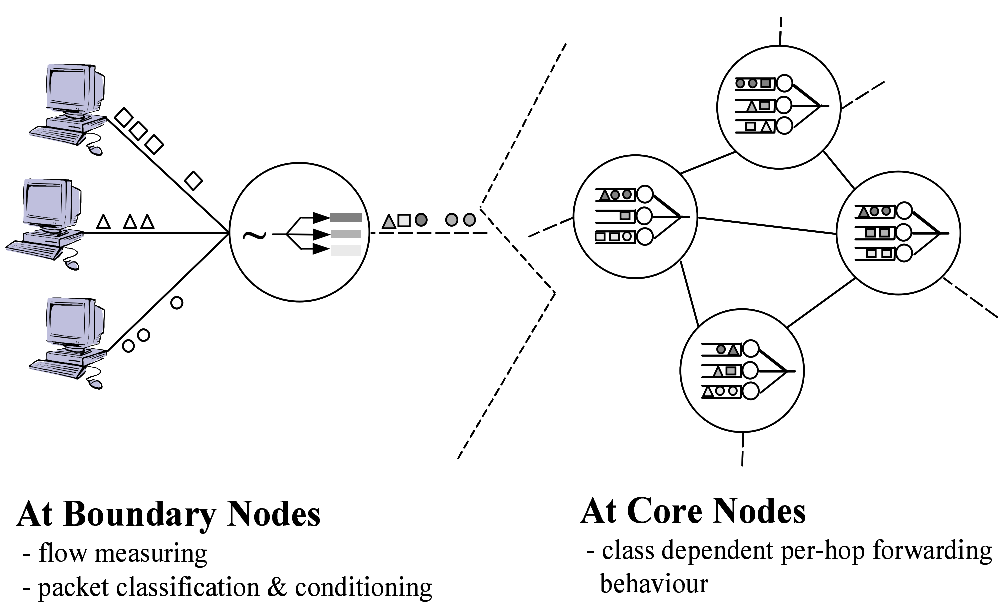



Figure 1: Illustration of the DiffServ Architecture and its Main Functions

nodes. Per-hop forwarding behaviors are implemented by means of buffer management and packet scheduling mechanisms in the core nodes. The observable packet forwarding behavior will often depend heavily on the relative load on the link, however, when several PHB aggregates are competing for buffer and bandwidth resources with different priorities, useful behavioral distinctions between PHB classes will be achieved.

The DiffServ architecture defines general Per-Hop Behavior group specification guidelines, but leaves the detailed implementation, i.e. packet classification and metering functions and PHB classes open for suggestions. Several PHB group approaches have been introduced, e.g., Assured Forwarding (AF) in [5] and Expedited Forwarding (EF) in [6]. In the following chapter the AF PHB group is described in more detail as further analysis in this report is mainly based on the AF PHB group.

## 2.1 Assured Forwarding PHB Group

Assured Forwarding PHB proposal defines a group of four independently forwarded PHB classes. Within each AF class, packet can be assigned to one of three different levels of drop precedence. The structure of the AF PHB group is shown in Figure 2. In case of long-term congestion, the drop precedence level defines the relative importance of the packet within the AF class. High drop precedence means that packet will be discarded more preferably compared to a packet with low drop precedence.

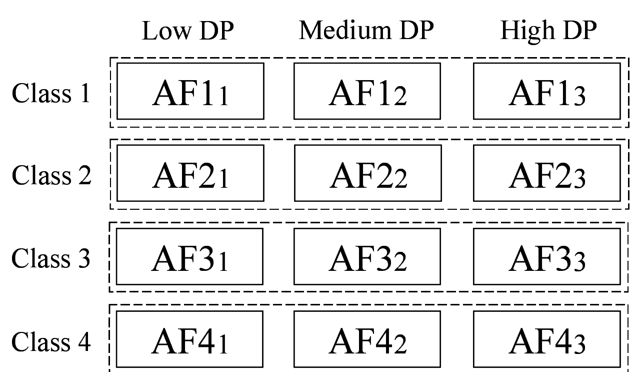



Figure 2: Structure of the AF PHB Group

The AF PHB group definition in [5] contains following specifications, which an AF PHB compatible DiffServ node must conform exactly for full compliance. Specifications define requirements for packet forwarding behavior and mutual relation of AF classes and drop precedence levels.

- *Packets in one AF class must be forwarded independently from packets in another AF class.*
- *A DiffServ node must not aggregate two or more AF classes together.*
- *A DiffServ node must allocate a configurable, minimum amount of forwarding resources (buffer size and bandwidth) to each implemented AF class.*
- *An AF implementation must specify how the excess resources are allocated between AF classes.*
- *Within an AF class, a DiffServ node must not forward an IP packet with smaller probability if it contains a drop precedence value $p$ than if it contains a drop precedence value $q$ when $p < q$.*
- *Within each AF class, a DiffServ node must accept all three drop precedence codepoints and they must yield at least two different levels of loss probability. If only two different drop precedence levels are implemented the codepoints AFx2 and AFx3 shall be combined to the higher drop precedence level.*
- *A DiffServ node must not reorder AF packets of the same microflow when they belong to the same AF class regardless of their drop precedence.*
- *An AF implementation must attempt to minimize long-term congestion within each class by dropping packets, while handling short-term congestion by queueing packets.*

The AF PHB can be implemented with a simple buffer mechanism. Each AF class has its own separated buffer with predefined buffer size and bandwidth allocation. Packets are divided into the buffers based on the AF class mark in the packet. The buffering mechanism will handle short-term congestion, i.e., packet bursts, by queueing packets into the buffer. Long-term congestion will be controlled by dropping incoming packets. The idea in this is to give equal treatment to flows that have different short-term properties but similar long-term characteristics. The dropping algorithm should compute smoothed buffer load to track long-term congestion properties of the buffer. This smoothed congestion level is then used in determining when packets should be discarded.

One approach in implementing the dropping algorithm is to use hard thresholds. Each drop precedence has a certain congestion level threshold. The packets aggregated to this drop precedence will be dropped, if the smoothed congestion level exceeds the threshold. Random Early Detection (RED) mechanism [4] can also be used in the dropping algorithm. The RED mechanism was invented to reduce transfer rate oscillation of TCP flows and thus improve throughput of TCP flows on congested nodes. The functioning of the RED mechanism is explained in Figure 3. The RED mechanism starts to drop incoming packets randomly with a certain probability after the congestion level reaches the threshold *th_min*. The dropping probability will increase linearly from zero to *p_max* until the congestion level limit *th_max* is reached, after which all the packets are dropped.

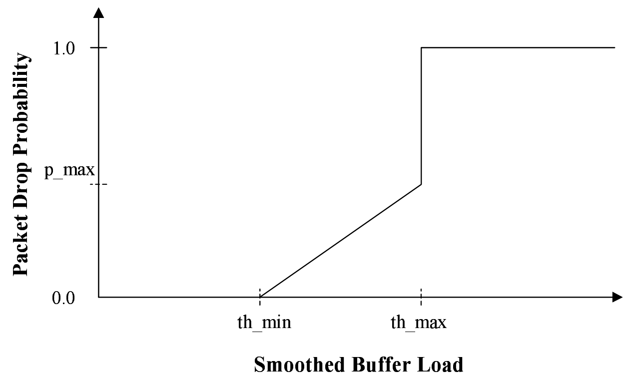



Figure 3: Illustration of the RED mechanism

The drop precedence levels may be implemented with hard thresholds, RED regions or combinations of these two. An example of AF PHB mechanism with four queues and combination of hard threshold and RED dropping mechanism is illustrated in Figure 4.

AF PHB definition in [5] does not specify any detailed parametrization for the PHB classes nor a mechanism for the packet classification and drop precedence setting. Packet classification mechanism could be, for example, so called Olympic service, which consists of three service classes, bronze, silver, and gold. Packets assigned to gold class would experience lighter load than the packets assigned to the silver class. Similar kind of relationship would exist between the silver and bronze classes. The olympic classes could be mapped to AF classes 1, 2 and 3. Inside a class, packets could be further separated by giving them either low, medium, or high drop precedence. The drop precedence can be assigned to a packet with, e.g., leaky bucket scheme. In the leaky bucket mechanism the packets are buffered in a virtual queue that is served with a predefined service rate. The drop precedence is assigned to arriving packet based on the occupancy level of the buffer, e.g., if queue occupancy level is high, the packet is marked to high drop precedence.

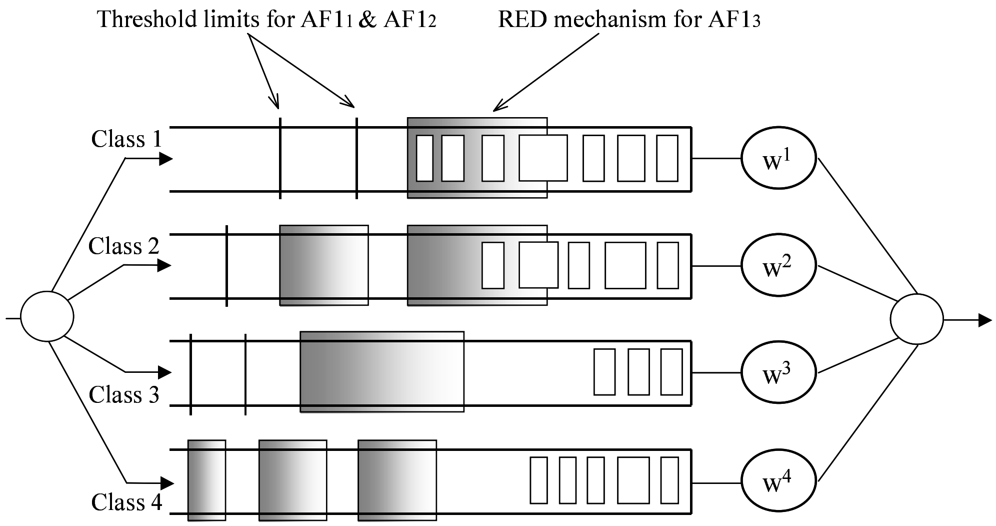



Figure 4: Example of AF PHB Implementation (Adapted from [7])

# 3 A Model of an AF PHB Node in a DiffServ Network

In this section an analytical model for an AF PHB node in DiffServ network with TCP and UDP flows is constructed. The model describes the functioning of the AF PHB mechanism in a single link network with one AF PHB node as depicted in Figure 5. It is assumed that several TCP and UDP sources send data into network. The data flows are divided into groups having the same transfer protocol (TCP or UDP) and the same traffic profile. In this context, traffic profile can be understood as a parametrization of packet classifier that determines how the packets are classified into the AF classes and the drop precedence levels. Thus, the flows within a flow group should experience similar forwarding behavior in the AF PHB node. The model is used to compute steady state throughputs of the flows in the AF PHB node. These results can be used to analyze fairness issues of the AF mechanism, e.g., how the TCP flow throughput relates to the predefined traffic profile and how the AF link bandwidth is allocated between different flows. Also an interesting issue is how the AF PHB mechanism can protect congestion sensitive TCP sources from greedy UDP sources that do not react to network congestion.

The modelling approach chosen is similar to the one presented by Nyberg et.al. in [10]. The model structure can be divided into two interdependent

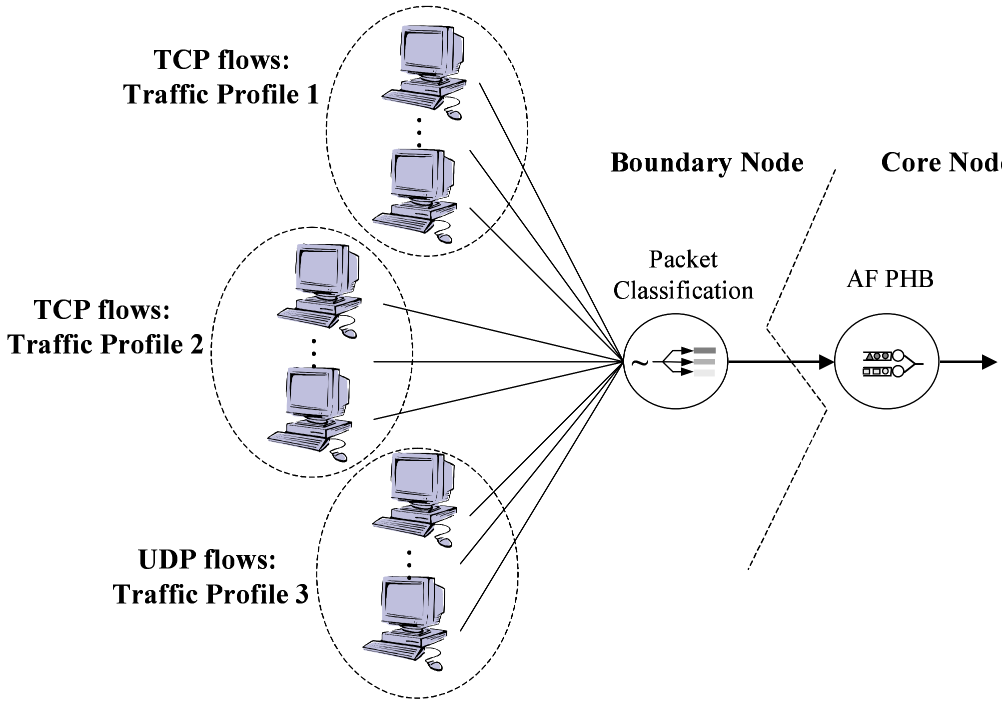



Figure 5: Modelled Network: Simple DiffServ Network with one AF PHB Node and Several Traffic Source Groups

submodels that are considered here separately. The first submodel describes the modelling efforts on the flow level, e.g., the TCP and UDP traffic, and packet classification mechanisms that are based on flow characteristics and traffic profiles. The other submodel is the AF PHB buffer model, which describes the packet level mechanisms of the AF PHB node. First in section 3.1 a model for both the TCP and UDP flows and the packet classification mechanisms are explained. In section 3.2 AF PHB buffer modelling with Markov chains is presented. Finally the interdependence of the submodels and overall structure of the model is explained in section 3.3.

## 3.1 Modelling Flow Level Operations

In this section, submodels for the traffic flows and packet classifier is presented. Traffic model describes the characteristics of the traffic flows that are generated from TCP and UDP sources. The packet classifier is modelled with a flow conditioner mechanism.

### 3.1.1 Traffic Model

Traffic flows are generated by the traffic sources according to TCP and UDP protocols. Further, it is assumed that the data packets within the flows are from Poisson process, i.e. the time between successive packets is exponentially distributed. Also the size of data packet is assumed to be exponentially distributed.

The TCP flows are divided into groups according to the traffic profile associated with the flow. Denote the set of TCP flow groups with $\mathcal{L}_{TCP}$. Respectively denote the set for UDP flow groups with $\mathcal{L}_{UDP}$. Let $\mathcal{L} = \mathcal{L}_{TCP} \bigcup \mathcal{L}_{UDP}$ denote the set of all flow groups $l \in \mathcal{L} = \{1, ..., L\}$. Each flow group consists of $n_l$ identical flows.

The UDP mechanism sends the data with constant intensity, and thus the UDP flows can be described with a traffic intensity value. The TCP mechanism is more complex. In congestion avoidance mode the TCP mechanism adjusts the sending rate according to a feedback signal from the network. The TCP congestion control gradually increases sending rate until a missing feedback signal indicates that a packet is lost due to network congestion. In case of packet loss, congestion control decreases the sending rate drasticly. After the response to the congestion, the TCP increases the sending rate until it again recieves an indication of congestion.

In this model, the TCP congestion control is assumed to follow differential equations for aggregates of TCP flows as described in [8]. To incorporate the TCP model with the AF PHB buffer model, it is assumed that the AF PHB buffer dynamics is faster than the dynamics of the TCP mechanism. This results in a stable point for TCP average sending rate $\lambda$ that depends on the round trip time (RTT) and the packet loss probability $p$

$$\lambda = \frac{1}{RTT}\sqrt{\frac{2(1-p)}{p}}. \tag{1}$$

Equation (1) relates the TCP congestion control with the AF PHB buffer via packet loss probability $p$. As the packet drop probability depends on the flow group $l$, the equation must be formulated for each TCP traffic profile group separately.

### 3.1.2 Packet Classifier Model

The packet classifier is implemented with a flow conditioner mechanism. Instead of tracing classifier mechanism on a packet level, traffic intensity (packet sending rate) characteristics of a flow are used in determining how the traffic is splitted among the drop precedence levels. The conditioner mechanism is based on the idea of splitting the flow across several levels of drop precedence. Two traffic intensity limits are used to split a flow into subflows that are marked with different drop precedence. These traffic intensity limits define the traffic profile associated with a flow.

In the following the conditioner mechanism is described in more detail. Consider a flow in group $l \in \mathcal{L}$. Let $\nu(l)$ denote the traffic intensity of flow $l$. Each flow is splitted into three different drop precedence levels $i \in \mathcal{I} = \{1,2,3\}$ with a flow conditioner mechanism. Traffic profile for a flow is determined with two traffic intensity limits $\nu_l^{\text{limit},1}$ and $\nu_l^{\text{limit},2}$, $0 < \nu_l^{\text{limit},1} < \nu_l^{\text{limit},2}$. Let $\lambda_l(i)$ denote the traffic intensity of the portion of flow $l$ that is marked with drop precedence $i$. Traffic profile and traffic intensity of the flow define $\lambda_l(i)$ by equations (2),(3) and (4).

$$\lambda_l(1) = \begin{cases} \nu(l) & ,0 \le \nu(l) \le \nu_l^{\text{limit},1} \\ \nu_l^{\text{limit},1} & ,\nu_l^{\text{limit},1} < \nu(l) \le \nu_l^{\text{limit},2} \\ \nu_l^{\text{limit},1} & ,\nu(l) > \nu_l^{\text{limit},2} \end{cases} \tag{2}$$

$$\lambda_l(2) = \begin{cases} 0 & ,0 \le \nu(l) \le \nu_l^{\text{limit},1} \\ \nu(l) - \nu_l^{\text{limit},1} & ,\nu_l^{\text{limit},1} < \nu(l) \le \nu_l^{\text{limit},2} \\ \nu_l^{\text{limit},2} - \nu_l^{\text{limit},1} & ,\nu(l) > \nu_l^{\text{limit},2} \end{cases} \tag{3}$$

$$\lambda_l(3) = \begin{cases} 0 & ,0 \le \nu(l) \le \nu_l^{\text{limit},1} \\ 0 & ,\nu_l^{\text{limit},1} < \nu(l) \le \nu_l^{\text{limit},2} \\ \nu(l) - \nu_l^{\text{limit},2} & ,\nu(l) > \nu_l^{\text{limit},2} \end{cases} \tag{4}$$

Traffic conditioner definition in equations (2),(3) and (4) split the flow into drop precedence levels such that $\nu(l) = \sum_{i=1}^{3} \lambda_l(i)$. Thus, percentual distribution of the flow into drop precedence levels can be expressed with

$$\delta_l(i) = \frac{\lambda_l(i)}{\nu(l)}, \forall i \in \mathcal{I}. \tag{5}$$

## 3.2 AF PHB Buffer Model

The model of the AF PHB buffer system is a simplified version of the original suggestion in [5]. Instead of four AF classes, two AF classes are implemented with two separated buffers. The packet dropping mechanism for AF buffer is implemented with hard thresholds and using instant value of buffer occupancy level. Assuming that packets arrive according to a Poisson process and the packet service times are exponentially distributed, the AF PHB buffer can be modelled as a queueing system with a Markov chain model, see, e.g.,

[2]. Using the Markov chain model, packet loss probabilities for each drop precedence aggregate can be computed. First in chapter 3.2.1 Markov model is constructed and analytic solution is presented for one buffer AF PHB node to clarify the modelling approach. In chapter 3.2.2 Markov model for AF PHB node with two dependent buffers is constructed.

### 3.2.1 One Buffer Case

A model for AF PHB node with one buffer is considered. Denote the acceptance thresholds for drop precedence level $i$ with $K_i$, $K_1 = K$, in which $K$ is the size of the buffer. Let $\lambda(i)$ denote the packet arrival rate into drop precedence class $i$ and $1/\mu$ mean value of the packet service time. Defining the cumulative sum of arrival intensities of drop precedence levels accepted into buffer as $\lambda_i = \sum_{k=1}^{i} \lambda(k)$, the buffer can be modelled as an $M/M/1/K$ queue on the state space $\{(m), 0 \leq m \leq K\}$, with state dependent arrival intensities as depicted in Figure 6.

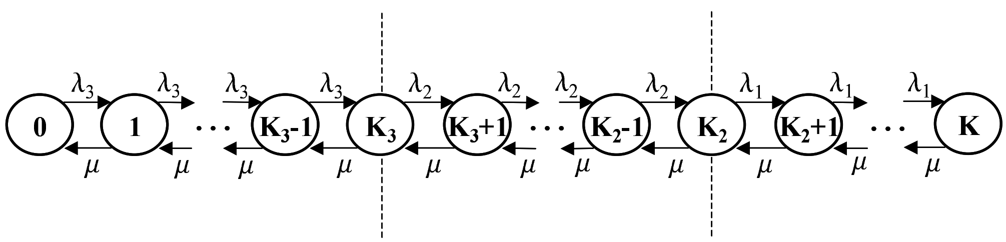



Figure 6: State Transition Diagram For One Buffer AF PHB Node

The stationary distribution of the buffer occupancy can be solved from the balance equations of the system, see, e.g., [2]. Let $\pi_m$ denote the equilibrium probability for state $m$. The global balance equations for the one buffer system can be written as

$$\begin{cases} \lambda_3 \pi_0 = \mu \pi_0 & ,m = 0 \\ (\lambda_3 + \mu)\pi_1 = \lambda_3 \pi_0 + \mu \pi_2 & ,m = 1 \\ \quad \dots & \dots \\ (\lambda_2 + \mu)\pi_{K_3} = \lambda_3 \pi_{K_3 - 1} + \mu \pi_{K_3 + 1} & ,m = K_3 \\ \quad \dots & \dots \\ (\lambda_1 + \mu)\pi_{K_2} = \lambda_2 \pi_{K_2 - 1} + \mu \pi_{K_2 + 1} & ,m = K_2 \\ \quad \dots & \dots \\ (\lambda_1 + \mu)\pi_{K-1} = \lambda_1 \pi_{K-2} + \mu \pi_K & ,m = K - 1 \\ \mu \pi_K = \lambda_1 \pi_{K-1} & ,m = K \end{cases} . \tag{6}$$

By solving the linear system of equations in (6), the $\pi_m$ can be defined for $m = 1, ..., K$ as a function of $\pi_0$,

$$\pi_m = \begin{cases} (\frac{\lambda_1}{\mu})^m \pi_0 & , m = 1, ..., K_3 - 1 \\ (\frac{\lambda_2}{\mu})^{(m-K_3)} (\frac{\lambda_1}{\mu})^{K_3} \pi_0 & , m = K_3, ..., K_2 - 1 \; . \\ (\frac{\lambda_3}{\mu})^{(m-K_2)} (\frac{\lambda_2}{\mu})^{(K_2-K_3)} (\frac{\lambda_1}{\mu})^{K_3} \pi_0 & , m = K_2, ..., K \end{cases} \qquad (7)$$

From the normalization condition $\sum_{m=0}^{K} \pi_m = 1$ the equilibrium probability $\pi_0$ for the empty buffer state can be solved.

As the buffer state probabilities $\pi_m$ are known, packet drop probability $p(i)$ for packets aggregated to drop precedence $i$ can be solved by summing up the probabilities of states $m \geq K_i$,

$$p(i) = \sum_{m=K_i}^{K} \pi_m. \qquad (8)$$

### 3.2.2 Two Buffers Case

AF PHB node with two buffers is modeled as two dependent $M/M/1/K$ queues as depicted in Figure 7. Mark buffers with $b \in \mathcal{B} = \{1, 2\}$. The packets in AF class 1 and AF class 2 are buffered in queue 1 and in queue 2, respectively. The buffers share the link bandwidth proportionally to weight parameters $w^b$, such that $w^1 + w^2 = 1$. Let $\mu$ denote the total service intensity. Thus, the service intensity for buffer $b$ is $\mu^b = w^b \mu$. If one buffer is empty, the other buffer gets all the bandwidth. Mark the acceptance thresholds for drop

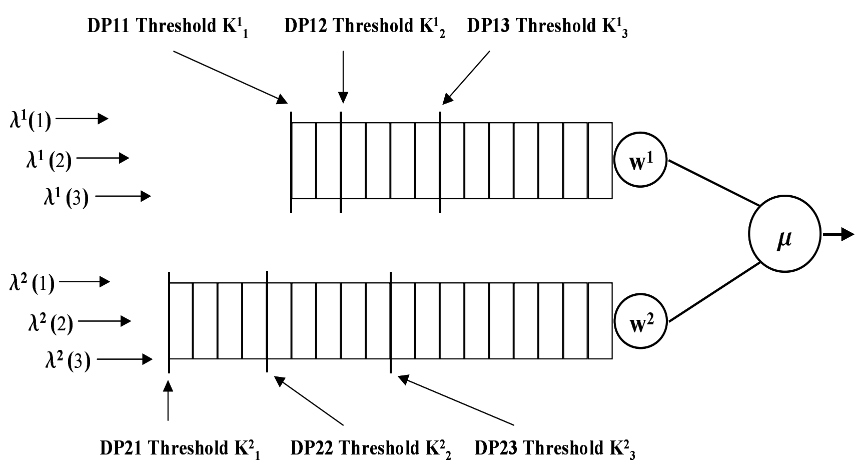



Figure 7: AF PHB Node With Two Buffers

precedence level $i$ in buffer $b$ with $K_i^b$, $K_1^b = K^b$, in which $K^b$ is the size of the buffer $b$. Let $\lambda^b(i)$ denote the packet arrival rate into drop precedence class $i$ in buffer $b$. Accumulated arrival intensities into drop precedence levels accepted into buffer $b$ is $\lambda_i^b = \sum_{k=1}^{i} \lambda^b(k)$. The two buffer system can be presented on the state space $\{(m^1, m^2), 0 \leq m^1 \leq K^1, 0 \leq m^2 \leq K^2\}$ as illustrated in the transition state diagram in Figure 8.

The stationary distribution of the two buffer system can be obtained in a similar way as in the one buffer case. In the model, the balance equations for the two buffer case are solved numerically to obtain the equilibrium probabilities for the buffer state $\pi_{(m^1,m^2)}$. Packet drop probabilities $p^b(i)$ for packets aggregated to drop precedence $i$ in buffer $b$ are

$$p^1(i) = \sum_{m^1=K_i^1}^{K^1} \sum_{m^2=0}^{K^2} \pi_{(m^1,m^2)}, \qquad (9)$$

$$p^2(i) = \sum_{m^2=K_i^2}^{K^2} \sum_{m^1=0}^{K^1} \pi_{(m^1,m^2)}. \qquad (10)$$

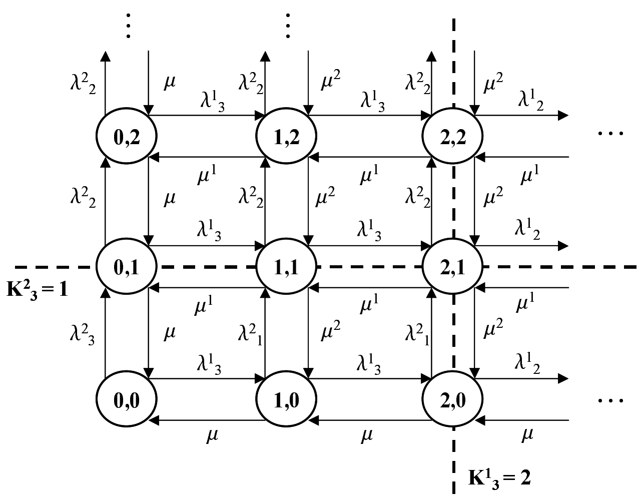



Figure 8: Example of State Transition Diagram for Two Buffer AF PHB Node

## 3.3 Top Level Structure of The Model

The top level structure of the model is presented here. The sending rate and traffic profile associated to the flow together with the flow conditioner mechanism determine how a flow is splitted into different drop precedence levels. The AF PHB buffer model is used to calculate the packet drop probability for the packets that are associated with a certain drop precedence. Based on the packet dropping probabilities for drop precedence levels, the packet loss probability of a flow can be computed. TCP congestion control adjusts the sending rate according to the feedback signal from the network, which in this case is represented by the packet loss probability at a flow. The model searches for a stable point, in which the TCP sending rate is in agreement with the achieved packet loss probability. If UDP flows are present, they increase traffic load into the AF PHB node and affect the packet drop probabilities and thus, the TCP congestion control mechanism.

The model top level structure is shown in Figure 9. Let $\nu(l)$ denote the sending rate of the flow associated with group $l$. Assume that flows in group $l$ are marked with AF class $b_l, b_l \in \mathcal{B}$ and thus they are buffered in queue $b_l$ in AF PHB node. The flow conditioner split the flow intensity into drop precedence levels as defined in equations (2),(3) and (4). Using the definition in equation (5), total traffic intensity of drop precedence level $i$ in buffer $b$ is thus

$$\lambda^b(i) = \sum_{l:b_l=b} n_l \delta_l(i) \nu(l). \tag{11}$$

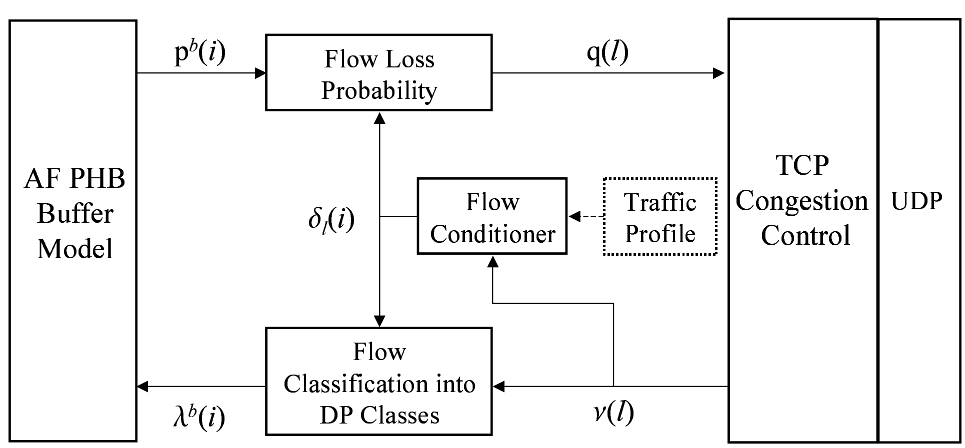



Figure 9: Schematic View of AF PHB Model With TCP and UDP Flows

Packet drop probabilities $p^b(i)$ for drop precedence levels can be determined from equations (9) and (10). The packet loss probability $q(l)$ of a flow is defined

$$q(l) = \sum_{i=1}^{3} \delta_l(i) p^{b_l}(i). \tag{12}$$

For each TCP group the sending rate is defined by equation (1) that relates the packet drop probability with the TCP flow average sending rate.

$$\nu(l) = \frac{1}{RTT}\sqrt{\frac{2(1-q(l))}{q(l)}}, l \in \mathcal{L}_{TCP}. \tag{13}$$

The sending rate $\nu(l)$ for UDP flows is constant and independent of the packet drop probability. Further, flow throughput rate can be defined as

$$\nu_{eff}(l) = \nu(l)\left(1 - q(l)\right), l \in \mathcal{L}. \tag{14}$$

The sending rate $\nu(l)$ of TCP flows in equilibrium point can be solved by solving the set of equations (5),(9)-(13) for the unknowns $\nu(l)$, $\lambda^b(i)$, $p^b(i)$, $\delta_l(i)$ and $q(l)$, where $i \in \mathcal{I}$, $l \in \mathcal{L}$, $b \in \mathcal{B}$.

# 4 Numerical Results

In this section some numerical results achieved with the AF model are presented. Various scenarios are analyzed both in one queue and two queue buffer systems to give insight how bandwidth is allocated among the flow groups and how the traffic load in the network affects the bandwidth allocation.

The parameters for the flow conditioner mechanism in equations (2)- (4) are adapted from the conditioner mechanism used in SIMA as described in [10]. This approach is chosen, because the SIMA conditioner mechanism utilizes only one input parameter, nominal bit rate (NBR). The NBR value defines how the network capacity is divided among different connections during overload situations and forms the basis of charging. Furthermore, the results gained from the AF PHB model are more easy to compare with the SIMA PHB model and can be used in future research efforts to evaluate the differences between AF and SIMA mechanisms.

In [10], priority level for a flow in group $l$ is determined by flow sending rate $\nu(l)$ and the NBR value, $nbr_l$. When the ratio of $\nu(l)$ to $nbr_l$ is one, the flow is given a medium priority. As the ratio doubles (halves) the priority level decreases (increases) by one unit. Thus, the $\nu(l)$ values, in which the priority class changes, can be determined by equation

$$limit_i(l) = nbr_l \cdot 2^{i-I/2-0.5}, i = \{1, ..., I-1\}, \tag{15}$$

in which the $I$ is a number of priority levels with $I$ implying the priority level with lowest drop precedence.

For the AF PHB, number of priority classes, i.e. drop precedence classes, is $I = 3$. Substituting $I$ into equation (15), the flow conditioner's intensity limits $\nu_l^{\text{limit},1}$ and $\nu_l^{\text{limit},2}$ can be directly solved,

$$\nu_l^{\text{limit},1} = nbr_l \cdot \frac{1}{2}, \tag{16}$$

$$\nu_l^{\text{limit},2} = nbr_l. \tag{17}$$

In the case of the flow conditioner mechanism described in the equations (2)-(4), $nbr_l$ value determines the traffic profile such that, if the sending rate $\nu(l)$ is considerably greater than the $nbr_l$ value, larger portion of the flow is associated with higher drop precedence and vice versa, if $\nu(l)$ is smaller than $nbr_l$, most of the flow is associated with low drop precedence.

## 4.1 One Buffer Case

The AF PHB model was solved in one buffer case with two TCP flow groups, $L = 2$. The model is used in evaluating how the differentiation in bandwidth allocation between the TCP flow groups is achieved with the AF PHB mechanism. The buffer size is set to $K = 39$ and the drop precedence limits for the buffer are $K_3 = 22$, $K_2 = 33$ and $K_1 = 39$. The buffer service intensity $\mu = 1$ and the round trip time $RTT = 1000$. The traffic profiles for the flow groups are $nbr_1 = 0.02$ and $nbr_2 = 0.04$.

Figure 10 shows the ratios $\nu_{eff}(1)/nbr_1$ and $\nu_{eff}(2)/nbr_2$ as a function of the network load, i.e. the number of active TCP sources. White color indicates that the flow throughput rate is equal to or greater than the NBR value, whereas darker tones indicate smaller throughput rate than the NBR value. Maximum and minimum value of the ratio are also presented in the figures. Figure 10 shows that the TCP flows in group 1 can achieve the desired throughput rate in a more congested network than the flows in group 2. If the network congestion level is low, both flow groups get more bandwidth than their NBR value indicate.

Figure 11 shows the bandwidth allocation between the flow groups. The figure represents the ratio $\nu_{eff}(2)/\nu_{eff}(1)$ as a function of the network load.

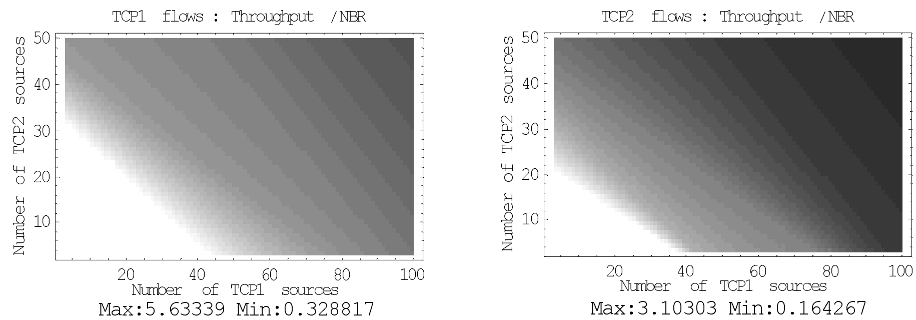



Figure 10: Throughput Rate for TCP 1 Flows (Left) and TCP 2 Flows (Right) as a Function of Number of Active TCP Sources $n_1$ (x-axis) and $n_2$ (y-axis).

The black areas correspond to equal bandwidth allocation, whereas the white areas correspond bandwidth allocation equal to ratio $nbr(2)/nbr(1) = 2$. Fixed bandwidth allocation ratio between the flow groups cannot be achieved, but the ratio depends on the network load. However, two regions with differentiated bandwidth allocation can be identified from the figure, although the ratio does not reach the value $nbr(2)/nbr(1)$. Differentiation in bandwidth allocation is not achieved, if the network load is considerably low or high.

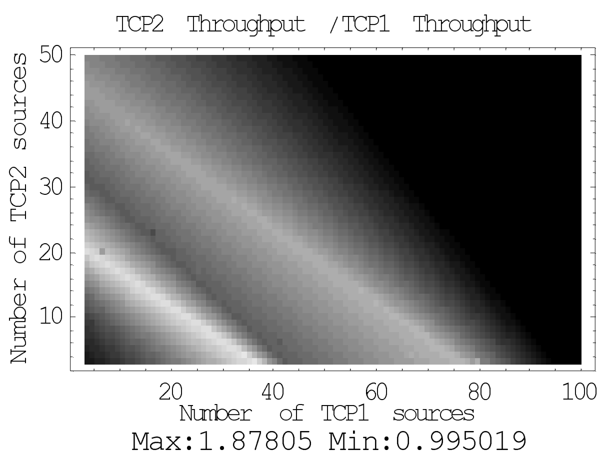
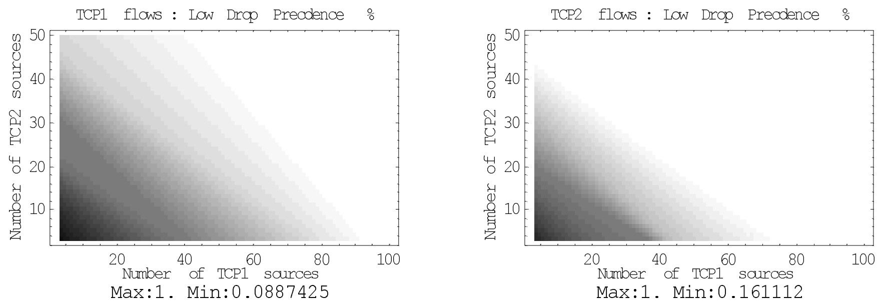
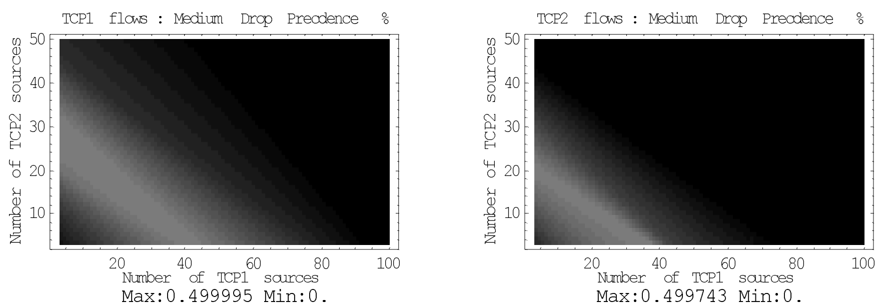



Figure 11: Bandwidth Allocation Between the TCP Flows as a Function of Number of Active TCP Sources $n_1$ (x-axis) and $n_2$ (y-axis)

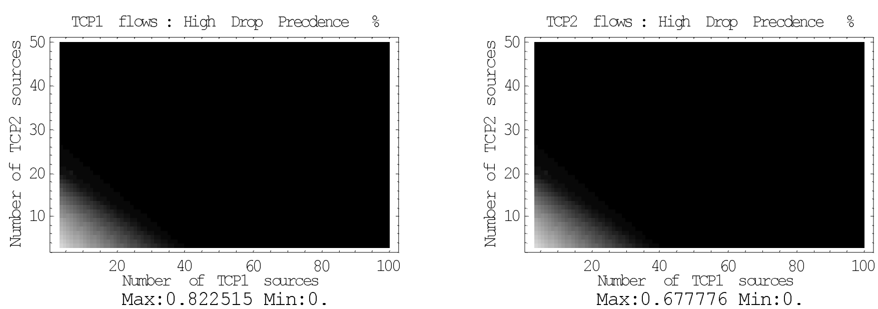
The differentiation in bandwidth allocation can be explained by considering how the flow conditioner splits the flow into the drop precedence levels. Figure 12 depicts the ratio $\delta_l(i)$ as a function of the network load in different scenarios with drop precedence levels $i = \{1, 2, 3\}$ and flows $l = \{1, 2\}$. $\delta_l(i)$ describes the proportion of flow $l$ that belongs to drop precedence level




(a) Low Drop Precedence, TCP 1 Flows (Left), TCP 2 Flows (Right).




(b) Medium Drop Precedence, TCP 1 Flows (Left), TCP 2 Flows (Right).




(c) High Drop Precedence, TCP 1 Flows (Left), TCP 2 Flows (Right).

Figure 12: Drop Precedence Allocation for the TCP Flows as a Function of Number of Active TCP Sources $n_1$ (x-axis) and $n_2$ (y-axis).

$i$ as defined in equation (5). In the figure, white areas correspond to $\delta_l(i)$ = 1, whereas the black areas correspond to $\delta_l(i) = 0$. When the network load is low, a flow can achieve a higher sending rate than the NBR value, and thus most of the flow is associated with high drop precedence as shown in Figure 12(c). When the network becomes more congested, the sending rate decreases, and the flows are splitted more evenly into the drop precedence levels. Eventually, when the network congestion level is high and the flow sending rate is decreased below the $\nu_l^{\text{limit},1}$ limit, the whole flow is set to low drop precedence. Thus, when network load is considerably low or high, most of both flows are marked into the same drop precedence level and differentiation in bandwidth allocation cannot be achieved.

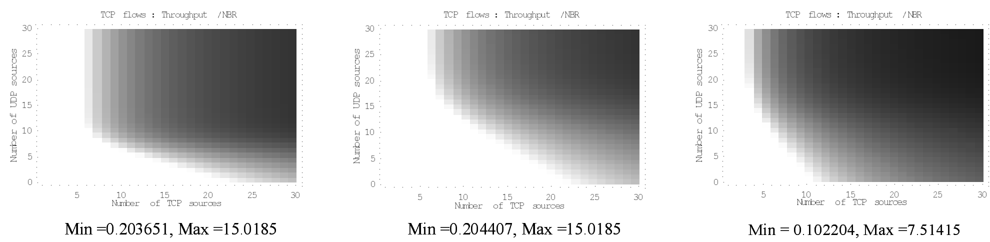
## 4.2 Two Buffers Case

The AF PHB model was solved for the two buffers case with TCP and UDP flow groups. The TCP and UDP flows are separated into different AF classes. The AF class 1 is reserved for UDP flows and handled in the first buffer. The buffer size $K^1 = 13$ and the drop precedence limits for the buffer are $K_3^1 = 7$, $K_2^1 = 11$ and $K_1^1 = 13$. The AF class 2 is reserved for TCP flows and handled in the second buffer. The buffer size $K^2 = 39$ and the drop precedence limits $K_3^2 = 22$, $K_2^2 = 33$ and $K_1^2 = 39$. Total service intensity $\mu = 1$, which is allocated to the buffers according to the weight parameters $w^1$ and $w^2$. The round trip time $RTT = 1000$ is used. For the UDP flows, the sending rate is set equal to the NBR value. Thus, half of the UDP flow's traffic intensity is assigned with low drop precedence and the other half with medium drop precedence.

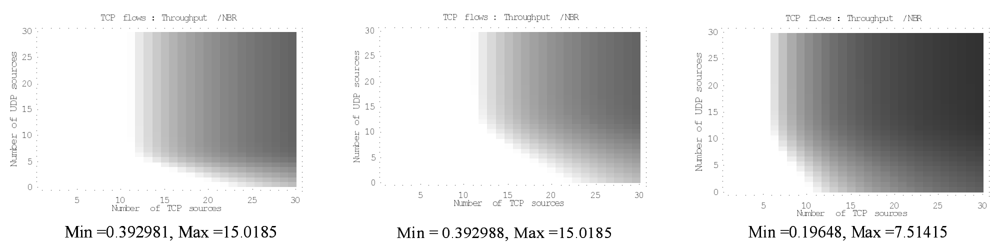
### 4.2.1 One TCP Group and One UDP Group

Using one TCP flow group, $l = 1$, and one UDP flow group, $l = 2$, it is studied how the AF PHB can protect congestion sensitive TCP flows from the greedy UDP flows. Also the effect of both the nominal bit rates, $nbr_1$ and $nbr_2$, and weight parameters, $w^1$ and $w^2$, is analyzed. Because the TCP flows are assigned to AF class 2, $b_1 = 2$, and UDP flows into AF class 1, $b_2 = 1$.

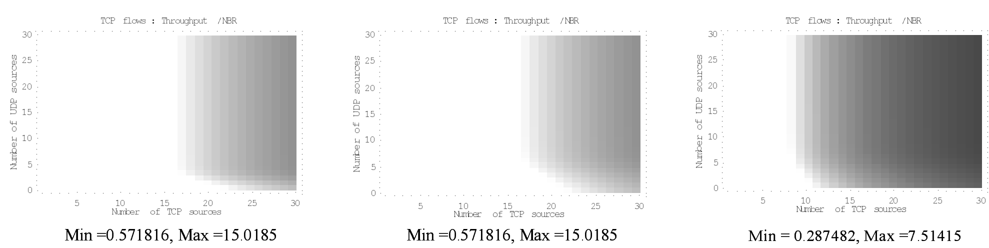
The TCP flow throughput rate can be described with ratio $\nu_{eff}(1)/nbr_1$. The ratio is solved as a function of the network load. The ratio is computed in different scenarios for weight parameter variations $w^1 > w^2$, $w^1 = w^2$, $w^1 < w^2$ and for nominal bit rates $nbr_2 > nbr_1$, $nbr_2 = nbr_1$ and $nbr_2 < nbr_1$ as shown in Figure 13. In white areas $\nu_{eff}(1)$ is equal to or greater than the $nbr_1$ value, while in the darker areas $\nu_{eff}(1)$ is smaller than $nbr_1$. The figure shows that a certain number of TCP flows can reach throughput rate equal to the NBR value independent of the number of UDP flows. Logically, as the weight parameter $w^2$ for the TCP buffer increases, more TCP flows can achieve the throughput rate equal to the NBR value. Also, if traffic load of the UDP flows in the network is small, the TCP flows can exploit the excessive bandwidth. The difference in the $nbr_1$ and $nbr_2$ values determines the shape of this transient area.




(a) $\{w^1, w^2\} = \{3/4, 1/4\}$; $\{nbr_1, nbr_2\} = \{0.04, 0.08\}$ (Left), $\{0.04, 0.04\}$ (Middle), $\{0.08, 0.04\}$ (Right).



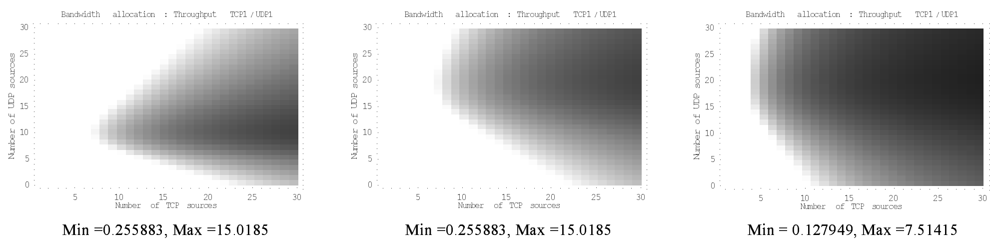

(b) $\{w^1, w^2\} = \{1/2, 1/2\}$; $\{nbr_1, nbr_2\} = \{0.04, 0.08\}$ (Left), $\{0.04, 0.04\}$ (Middle), $\{0.08, 0.04\}$ (Right).



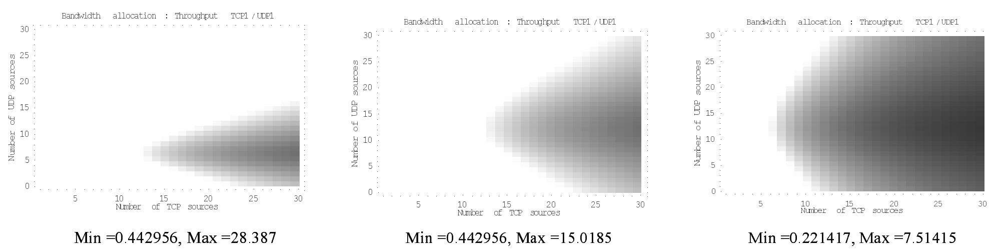

(c) $\{w^1, w^2\} = \{1/4, 3/4\}$; $\{nbr_1, nbr_2\} = \{0.04, 0.08\}$ (Left), $\{0.04, 0.04\}$ (Middle), $\{0.08, 0.04\}$ (Right).

Figure 13: Throughput Rate for TCP Flows as a Function of Number of Active TCP Sources $n_1$ (x-axis) and Active UDP Sources $n_2$ (y-axis). Scenarios Show the Effect of Variations in Parameters $nbr_l$ and $w_i$.

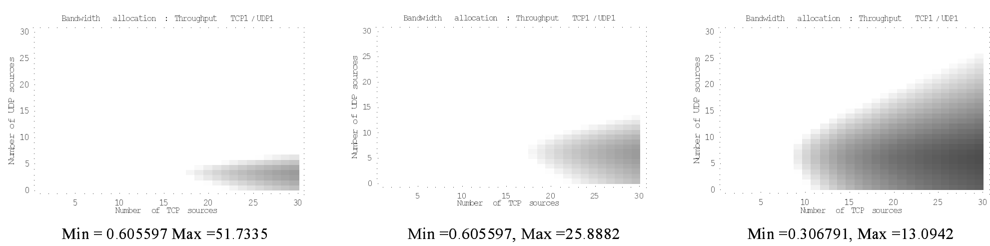
The bandwidth allocation between the TCP and UDP flows can be described with the ratio $\nu_{eff}(1)/\nu_{eff}(2)$. In Figure 14 the ratio is solved as a function of the network load with the same variations of weight parameters and NBR values as in Figure 13. The black areas correspond to the bandwidth allocation ratio less than the ratio $nbr(1)/nbr(2)$, whereas the white areas correspond to the bandwidth allocation ratio equal to or greater than




(a) $\{w^1, w^2\} = \{3/4, 1/4\}$; $\{nbr_1, nbr_2\}$= {0.04, 0.08} (Left), {0.04, 0.04} (Middle), {0.08, 0.04} (Right).




(b) $\{w^1, w^2\} = \{1/2, 1/2\}$; $\{nbr_1, nbr_2\}$= {0.04, 0.08} (Left), {0.04, 0.04} (Middle), {0.08, 0.04} (Right).




(c) $\{w^1, w^2\} = \{1/4, 3/4\}$; $\{nbr_1, nbr_2\}$ = {0.04, 0.08} (Left), {0.04, 0.04} (Middle), {0.08, 0.04} (Right).

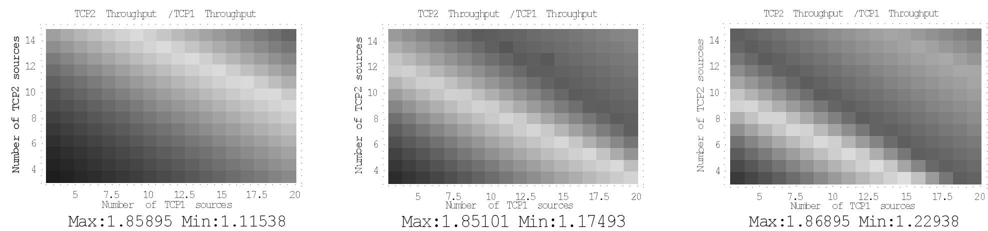
Figure 14: Bandwidth Allocation Between TCP and UDP Flows as a Function of Number of Active TCP Sources $n_1$ (x-axis) and Active UDP Sources $n_2$ (y-axis). Scenarios Show the Effect of Variations in Parameters $nbr_l$ and $w_i$.

the ratio $nbr(1)/nbr(2)$. The figure depicts the situation from the viewpoint of the TCP flows, because in the white areas the UDP flows may get less bandwidth than in the proportion to the NBRs. As the number of UDP flows increases, the TCP flows cannot exploit the bandwidth allocated for the UDP buffer anymore. This leads to decrease in the minumum number of TCP flows that are able to get bandwidth at least in the proportion to the NBRs. However, as the number of UDP flows increases over a certain limit, the minumum number of TCP flows starts to increase. This effect is due to the UDP flows, whose throughput rate is decreasing as more UDP flows compete from the bandwidth reserved for the UDP buffer. The effect of the weight parameters and the NBR values on the limiting value of UDP flows can be analyzed by comparing Figures 14(a), 14(b) and 14(c). The limiting value for the number of UDP flows increases as more bandwidth is reserved for the UDP buffer or if the NBR value of the UDP flows decreases.

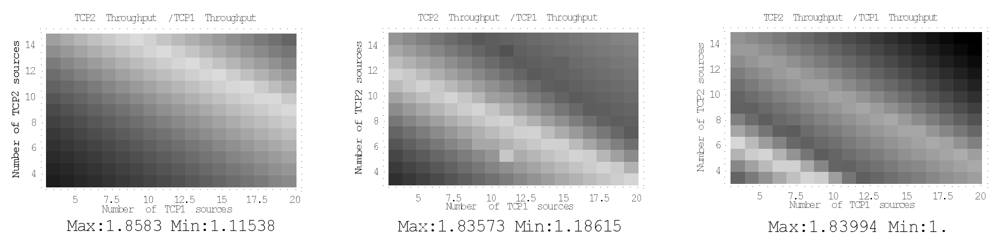
### 4.2.2 Two TCP Groups and One UDP Group

Using two TCP flow groups, $l$=1, $l$=2, and one UDP flow group, $l$=3, the effect of UDP flows on the bandwidth allocation between the TCP flow groups is studied. The TCP flows are assigned to AF class 2, $b_1 = 2$ and $b_2 = 2$, whereas the UDP flows into AF class 1, $b_3 = 1$. The nominal bit rates of the flow groups are set to $nbr_1 = 0.02$, $nbr_2 = 0.04$ and $nbr_3 = 0.08$.

Bandwidth allocation between the TCP flow groups is described with ratio $\nu_{eff}(2)/\nu_{eff}(1)$. The ratio is solved as a function of TCP traffic load in the network. The ratio is computed in six cases with different number of active UDP sources $n_3$ and weight parameter variations $w^1 = w^2$ and $w^1 > w^2$, as shown in Figure 15. As in Figure 11, the black areas correspond to equal bandwidth allocation, whereas the white areas correspond bandwidth allocation equal to ratio $nbr_2/nbr_1 = 2$. The figure indicates that the UDP flows have an effect on the bandwidth allocation between the TCP flow groups. If the number of active UDP sources increases, the differentiated area shrinks towards the lower TCP traffic load. This effect is likely due to bandwidth sharing between the buffers, which enables the TCP buffer to get more bandwidth if the traffic load into UDP buffer is low. The bandwidth allocation between the TCP flows eventually becomes insensitive to UDP traffic if the number of UDP flows $n_3$ increases over a certain limit. In this case, the limiting number of UDP flows is 10. At the saturation point, the bandwidth allocation level depends on the weight parameters $w^1$ and $w^2$, as can be recognized by comparing the right most scenarios in Figures 15(a) and 15(b).




(a) $\{w^1, w^2\} = \{1/2,\ 1/2\}$; $n_3 = 1$ (Left), 4 (Middle), 10 (Right).




(b) $\{w^1, w^2\} = \{2/3,\ 1/3\}$; $n_3 = 1$ (Left), 4 (Middle), 10 (Right).

Figure 15: Bandwidth Allocation Between TCP 2 and TCP 1 Flows as Function of Number of Active TCP Sources $n_1$ (x-axis) and $n_2$ (y-axis). Scenarios Show the Effect of Variations in the Number of UDP Flows $n_3$ and in the Weight Parameters $w^1$ and $w^2$.

# 5 Conclusions

In this report, functioning of an AF PHB node in a DiffServ network with TCP and UDP traffic was analyzed. The AF PHB node supporting two AF classes with three drop precedence levels was modelled at packet level using a Markov chain model. The AF packet classification mechanism was implemented with a flow conditioner mechanism that is based on the flow's characteristics and a traffic profile associated with the flow. Congestion sensitive TCP flows are incorporated into the packet level model using a relation of the TCP stable sending rate and the packet loss probability. The model was used in studying how the differentiation in bandwidth allocation can be achieved with the AF PHB mechanism. It was also analyzed how the AF PHB mechanism can protect congestion sensitive TCP sources from greedy UDP sources.

Various scenarios were numerically solved to give insight how the bandwidth is allocated among the flow groups with similar traffic profile and how the network load affects the bandwidth allocation. Traffic profile associated with the flow was determined by the NBR value. Based on the numerical results, differentiated bandwidth allocation was achieved with the AF PHB mechanism. However, a fixed bandwidth allocation ratio between the flow groups was not achieved. Instead, the network load affects the ratio, which varies from equal bandwidth allocation to a slightly less than proportion of the purchased NBRs. When TCP and UDP flows are separated into different AF classes, i.e. the flows are handled in separate buffers, the AF PHB mechanism can protect TCP flows from the greedy UDP flows. The minimum guaranteed bandwidth for the TCP flows can be determined with buffer weight parameters. The differentiated bandwidth allocation between the TCP and UDP flows can be achieved. However, the ratio is not constant but depends on the network load. The variation in the UDP traffic load affects also the bandwidth allocation between the TCP flows, and thus, the TCP flows are not fully protected from the effects of the UDP flows.

The model introduced in this report describes a reduced version of the AF PHB mechanism. Further research is needed to develop a model that better represents the actual AF PHB mechanism. An essential improvement would be to extend the AF PHB buffer model to contain a packet dropping algorithm based on smoothed buffer congestion level. Also some other improvements for the AF PHB model should be carried out, e.g., the increase of the number of AF classes up to four to evaluate the value of using several AF classes. By discretization, the smoothed congestion level variable could be attached into the AF PHB Markov model. However, even with only one AF class, this approach may lead to a situation, which is computationally inefficient to solve. One direction in the development of the model is to enhance the modelling of the TCP flows. Currently the TCP model is based on long term behavior of the TCP flows. A more realistic approach should include also short term TCP connections, whose arrival and duration is based on a stochastic process. Furthermore, the model could be extended to cover a set of sequentially connected AF PHB nodes. This modelling approach would provide information about functioning of the AF PHB mechanism in a small network.

# References


[1] Blake S., Black D., Carlson M., Davies E., Wang Z. and Weiss W., *An Architecture for Differentiated Service*, December 1998, IETF RFC 2475.

[2] Daigle J. N., *Queueing Theory for Telecommunications*, Addison-Wesley, 1992.

[3] Di Z., Mouftah H. T., *Performance Evaluation of Per-Hop Forwarding Behaviours in the Diffserv Internet*, Proceedings of Fifth IEEE Symposium on Computers and Communications, 2000, pp. 334 -339.

[4] Floyd S. and Jacobson V., *Random Early Detection Gateways for Congestion Avoidance*, IEEE/ACM Transactions on Networking, 1993, vol 3., pp. 397-413.

[5] Heinänen J., Baker F., Weiss W. and Wroclawski, *Assured Forwarding PHB Group*, June 1999, IETF RFC 2597.

[6] Jacobson V., Nichols K. and Poduri K., *An Expedited Forwarding PHB*, June 1999, IETF RFC 2598.

[7] Kilkki K., *Differentiated Services for The Internet*, MacMillan Technical Publishing, 1999.

[8] Kelly F., *Mathematical Modelling of the Internet*, Proceedings of Fourth International Congress on Industrial and Applied Mathematics, 1999, pp. 105-116.

[9] May M., Bolot J.-C., Jean-Marie A., Diot C., *Simple Performance Models of Differentiated Services Schemes for the Internet*, Proceedings of IEEE INFOCOM'99, 1999, vol. 3, pp.1385 -1394.

[10] Nyberg E., Aalto S., Virtamo J., *Relating Flow Level Requirements to DiffServ Packet Level Mechanisms*, Submitted, 2001.

[11] Postel J., *Internet Protocol*, September 1981, IETF RFC 791.

[12] Wroclawski J., *The Use of RSVP with IETF Integrated Services*, September 1997, IETF RFC 2210.

[13] Vutukury S., Garcia-Luna-Aceves J.J., *A Scalable Architecture for Providing Deterministic Guarantees*, Proceedings of Eight International Conference on Computer Communications and Networks, 1999, pp. 91 -96.