Performance Analysis of a new Mobility/QoS-aware Architecture

Nuno Vasco Lopes*, Maria João Nicolau†, Alexandre Santos*
*Department of Informatics, †Department of Information Systems
Algoritmi Center
University of Minho, Braga 4710-057, Portugal
email: vascolopes@di.uminho.pt, joao@dsi.uminho.pt, alex@di.uminho.pt

***Abstract*—Ideally, the future Internet must provide acceptable Quality of Service (QoS) to mobile users that are running real-time applications and are moving across different access points at high speeds. The user mobility presents a great challenge to the network layer in order to maintain users on going connections. Currently, the Internet protocol that manages the user mobility at the network level is the Mobile Internet Protocol (MIP). This protocol, when a mobile user changes its point of attachment, maintains the same IP address for mobile node, so that user mobility became invisible to the application level and thus avoiding a connection interruption. Although, MIP standard allows the user mobility while maintaining an uninterrupted connection to an application, it does not have any concerns with the QoS support provided to applications with more strict performance requirements such as real-time applications. This paper addresses the issue of mobility and QoS management principles as well as the mobility and QoS management integration in the sense of build a QoS-aware architecture for mobile Internet. After covering the mobility and QoS management principles and integration, this paper also proposes a new QoS-aware architecture for mobile Internet. This new architecture takes into account the specific characteristics of mobile networks in order to design an integrated Mobility/QoS-aware management architecture suitable for real-time applications requirements. The simulation results indicate that the suggested architecture is able to provide acceptable QoS levels to real-time applications that are running in mobiles devices.**

***Index Terms*—Mobile Internet, Quality of Service, Mobility Management, QoS Management, QoS and mobility Integration**

I. Mobility and QoS Management: Principles

During the last years, several network communications challenges have arisen with a growing number of users demanding at first QoS, then mobility, and now the both simultaneously. The first challenge which aimed to answer those demands consisted in enabling the support of QoS to fixed Internet. The second challenge was to endow Internet with mobility support. Now days, the challenge is to adapt the existing QoS models to mobile Internet making them suitable for incoming Fourth Generation (4G) networks. Although mobile Internet has also introduced its own challenge of becoming effective for micro-mobility scenarios with frequent handovers. Therefore, the current challenge is much more than adapting the existing QoS models to mobile Internet in general. The challenge is to enhance the mobility management of Mobile IP for micro-mobility scenarios and simultaneously design a QoS solution suitable for all types of mobile environments (macro and micro). To enhance mobile protocol standard for the Internet the micro-mobility protocols have been proposed to overcome the unreasonably high signaling load and latency problems associated with Mobile IP in scenarios where the Mobile Node (MN) moves frequently within a single administrative domain, but by themselves they do not provide QoS support for real-time applications. And unfortunately, the existing QoS models were designed for fixed Internet thus, they do not take into account mobility when they perform resource management, resulting in unsuitable solutions for mobile Internet which are unable to deal with the changeable nature of wireless networks [21].

The wireless networks allows the Internet users to be mobility enable. The nodes inside a wireless networks do not need to be necessary mobile (e.g. in wireless home network a stationary printer) and on the other hand a MN do not always requires a wireless link (e.g. a professor that uses a wired notebook at the school, can at the end of the working day take home the notebook to work on it at night). There is different levels of user mobility, for instance if a user moves within the same subnet, this user will not even be mobile from the perspective of network layer. A user that moves between different subnets, but shutdowns the MN while moving between networks, is mobile from the perspective of network layer, but nevertheless the connections are not maintained during the mobility. Another level of user mobility, is what we call a full mobile user which takes its MN from one location to another (locations in different networks) maintaining an ongoing connection while moving between the networks. So, there are many forms of mobility, but perhaps the most interesting and challenge form of mobility is the one in which users are both wireless and mobile, and the connection to wireless link or links are maintained uninterrupted with mobility. For instance, a scenario where a user maintains a voice application over the Internet while driving on the motorway.

This type of mobility puts same challenges to the network layer such as locating a mobile user, routing the data packets to the mobile user and to handle with the changes of its point of attachment.

In order to maintain an ongoing connection the MN must also maintain the same IP address while is moving between the networks, because an Internet application uses an end-to-end logical link to communicate - service given by a transport protocol - which in turn is identified by an IP address and a port number. If the MN is able to maintain its IP address the mobility became transparent to application layer and the applications layer do not need to be concern if a user is mobile or not. This task of maintaining the same IP address when a MN changes to a new network is carry out at network layer by the MIP. For this purpose MIP defines two important entities: Home Agent (HA) and Foreign Agent (FA). This two entities perform the mobility management functions related to the MN mobility. In order for mobility user to be invisible to application layer, a MN has two IP addresses, one permanent given by the original network, i.e. Home Network (HN) and a temporary one known as Care-of-Address (CoA) which is given by the visited or Foreign Network (FN). The temporary IP address serves to the network layer knows the new localization of MN, and in this way it can route datagrams to mobile. A correspondent that wants to send a datagram only needs to addresses the datagram to the MN's permanent address, then it is up to HA and FA perform the mobility management functions needed to guarantee that the datagram arrives at the target MN. The agent in the HN is responsible to track the FN in which the MN is attached and

Nuno V. Lopes is supported by a FCT Grant (SFRH/BD/35245/2007)

the agent in the FN must advertise the HA that the MN resides in its network and inform that it was given a certain CoA. A second role of HA is to intercept the datagrams addressed to MN and forward them to the foreign agent, doing it with the encapsulation of the original datagrams within a new datagram addressed to the MN's CoA.

When a MN moves to a new location it locates the FAs through a process known as agent discovery. The agent discovery can be made either via agent advertisement or via agent solicitation. With the former, a foreign agent sends a advertisement message, using an extension of router discovery protocol (RFC 1256), for advertise its services to the MN. With the latter, the MN sends a broadcast solicitation message instead of waiting to receive an advertisement from a foreign agent, and an agent receiving this message responds with a unicast advertisement message directly to the MN.

After the MN has received the CoA, this address must be registered at the HA, either via foreign agent or directly by the MN.

As we know today's Internet provides a best-effort service to all types of applications. Independently of the strict delay requirements that a real-time application could have all the packets are treated equally at the routers. In the next paragraphs will be described the main QoS architectural components and the underlying principles which should be included in the Internet in order to real-time applications can have acceptable quality network services in scenarios of congestion.

In order for packets in a router's queue to be recognized as belonging to class of traffic with higher time constrains, each packet of that class must be marked. Having the packets organized in classes of traffic a router can treat the classes of traffic with a different policy. This is perhaps the first principle to provide QoS guarantees to applications with time or loss constrains.

Now, suppose that two applications with the same class of traffic start sending packets, but one of them is malicious and increases its data rate sending much more packets, leading to the starvation of the other application. Therefore, it is needed a mechanism that protects one flow from another misbehave flow. This mechanism has the job of police if the traffic flows meet certain criteria (for instance, if it not exceeds the peak rate of 1Mbps). If a flow misbehaves, the policing mechanism performs some action so that traffic be in conformation with the traffic profile. Therefore another principle will be the provision of isolation among flows.

Suppose now that the combined data rates of two flows belonging to the same class of traffic are beyond the link capacity, in such scenario each application will lose the surplus packets. If they were two audio applications the both would have an unacceptably QoS, thus completely unusable even for a little demanding user. Therefore, there is no advantage in allowing a flow into the network if there is not enough resources to accommodate the flow, because with the admission of flow the resources are being used to transmit packets of a flow that has no utility to the user. Therefore another principle, is the need of a call admission process in which applications require its QoS needs to the network and then they are admitted or blocked (if the required QoS can not be provided). With the call admission process will be an allocation enforcement of a certain amount of the link bandwidth that a flow can use during the transmission of its traffic, however this reserved bandwidth can not be utilized by other applications even if it is not being totally used. It is therefore also important and recommendable to make an efficient bandwidth utilization by enabling the others flows to use all the unused link bandwidth, i.e. that it is not being used by the existing flows. Thus, the last principle is to make an efficient resource management in order to maximize the link bandwidth utilization.

These principles are the main basic pillars in providing QoS guarantees for Internet applications. These principles should dictate the behaviour of a generic QoS architecture. These principles can help and guide the network designers in the development of their QoS frameworks.

Addressing the support of QoS for real-time applications in the IP networks as a global architecture implies a broad design process at different functional planes such as management, control and data planes. Where the data and control planes are responsible for providing the essential QoS mechanisms to control real-time traffic based on requested QoS made by applications whereas the management plane is responsible for ensuring that QoS commitments assumed by the network are assured by configuring network resources accordingly to those QoS commitments.

To endow the Internet with these principles several components can be used and implement in the Internet, for instance, scheduling and policing mechanisms, signaling protocols for call setup, admission control algorithms and so on.

II. Related Work

According to literature, there are several QoS architectures proposals for Mobile IP networks however, the research community has not yet decided which is the best solution. Some aspects of which are not consensual are: if the management plane is centralized or distributed, if the dissemination of control messages are in-band or out-of-band, the scalability in large scale scenarios, and the suitability to mobile networks among others.

On the other hand, most of the existing solutions for QoS provisioning within mobile environments are grounded on the deterministic service model. The use of a deterministic service model for QoS provision in high dynamic networks where mobile nodes are continuously changing their point of attachment has proved to have elevated costs in terms of state information maintenance, signaling, processing and consistency [3], [11], [10], [7], [5].

Therefore, there is an emerging demand for designing a QoS architecture suitable for those wireless environments. Some of the most relevant solutions that have been proposed in this area are the QoS-Conditionalized Handoff for Mobile IPv6, the RSVP extensions for Mobile IP and DiffServ extensions to Mobile IP.

The QoS-Conditionalized Handoff for Mobile IPv6 [8] scheme is built over the hierarchical mobile IPv6 in order to be suitable for micro-mobility scenarios. The main idea is to employ the QoS hop-by-hop option of the binding message of MIP in order to provide the QoS requirements of the mobile device to the routers along the new data path towards new access router. The routers based on the resource availability and on QoS requirements containing in the binding message decide if the handover is possible or not. Making QoS conditionalized handovers without the need of using a signaling protocol is a great advantage because reduces the amount control messages in transit within network, although has the disadvantage that all nodes must be modified in order to implement the required functionality.

The first RSVP extension proposal was the Mobile RSVP [2], a protocol that makes advanced reservations in multiple locations where an MN may possibly go. However, the solution has the problem of creating excessive resource reservations causing the waste of bandwidth and reducing the network performance.

The HMRSVP [19] solution, another RSVP extension, improves the MRSVP with local MN's registrations and advanced reservations only for inter-domain handovers. But still has a significant processing burden and resource waste, and is also restricted to HMIPv6 networks.

Another MRSVP derived solution is proposed in [20] where the authors introduce a Crossover Router (CR) entity to reduce tunnel distance between previous access router and new access router created by the FMIPv6 protocol. The CR is responsible for intercepting all packets sent to MN's previous CoA and forward them to the new access router. To deliver the QoS requests, they extend Fast Binding Update (FBU) and Handover Initiate (HI) messages, which are used for informing the new access router of the MN's QoS requirements. With the information of the MN's QoS requirements, the new access router can make an advanced reservation on the common data path. This solution is claimed to outperform MRSVP in terms of signaling cost, reservation re-establishment delay and bandwidth requirements. However, the solution introduces more signaling messages and complexity.

In [4] the authors deployed a modified RSVP called Mobility-Aware Resource Reservation Protocol (MARSVP) where the binding update and the binding acknowledgment messages are conveyed in two new RSVP objects, these new RSVP objects must be added to the standard RSVP messages. This solution implies modifications on MIPv6 and RSVP protocols, and on end nodes.

The DiffServ extensions to Mobile IP are mainly designed using a Policy-Based Management System (PBMS). A PBMS can be implemented in order to make the QoS management between adjacent DiffServ domains, as is the case of the work in [12] or between a DiffServ domain and a IntServ domain, as is the case of the work in [15].

The use of policy-based management systems such as a centralized Bandwidth Broker (BB) entity, for coordinating the network resources is in fact one more element to add to the QoS architecture, because the architecture still needs a QoS model and a signaling protocol to communicate the policy information. Furthermore, BBs are centralized resource management entities which have a complex implementation because they need to congregate several functionalities into a single entity. Moreover in high dynamic networks such as wireless networks, rather than being a solution they can became the network bottleneck [17].

In summary, offering a deterministic service model with strict guarantees as those proposed by the above solutions seems to be hardly possible due to the non-deterministic nature of mobile networks.

III. MODEL PROPOSAL

The goal is to build a model where the underlying Mobility/QoS-aware architecture is constructed with dynamic QoS functionalities, adaptive resource management and seamless handovers. Another major goal is to minimize the scalability problems such as signaling overhead, processing and state load that can arise in scenarios where handovers are more frequent. The model was designed in presumption that in wireless networks, the most critical points are the access routers on account of wireless link constraints. Thus, main concern of the proposed model is to quickly reestablish the QoS in these critical points whenever a handover occurs.

To deal with the inefficiency of MIPv6 in micro-mobility scenarios the model proposes a specific integration of two micro-mobility protocols: the Fast Mobile Internet Protocol version 6 (FMIPv6) [1] and the Hierarchical Mobile Internet Protocol version 6 HMIPv6 [18]. From now on this integration will be called the F-HMIPv6. The F-HMIPv6 improves the MIPv6 mobility management with seamless handovers and local handover registrations. The F-HMIPv6 reflects the recommendations defined in RFC 4140 except in proceedings of Handover Initiate (HI) and Handover Acknowledgment (HAck) messages which have been maintained between the previous and new access routers as in FMIPv6 standard.

In what respects to QoS architecture, the proposed model extends the Resource Management Function (RMF) of DiffServ in the edge routers with new functionalities. Basically, the new RMF consists of three components: QoS model - Diffserv QoS mechanisms to treat priority traffic differently; Admission Control - Admission control to determine whether or not a node has sufficient resources to support the requested QoS and; Dynamic Allocator - Reallocation mechanism that reallocates more bandwidth for handover flows belonging to priority classes. The resource allocation is only made in the edge routers; it is assumed that interior nodes are engineered and reconfigured taking into account their DiffServ characteristics, the routing behavior as well as the maximum aggregated traffic per-class injected into the domain through the ingress router.

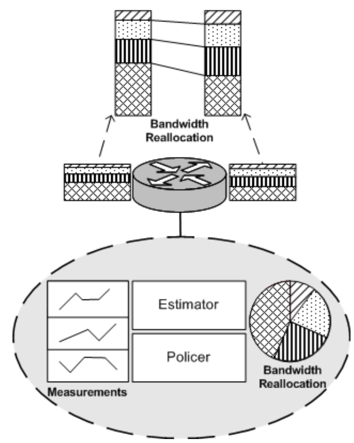


Figure 1: Resource Management Main Functions

Figure 1 illustrates the four main functions of the new RMF (Measure, Estimate, Police and Reallocate bandwidth). Estimators implement measurement mechanisms which have the responsibility of determine the current network load in each DiffServ class and the current load in use by a MN each DiffServ class (i.e. the MN's QoS Context).

The policer runs an algorithm that decides the admission or rejection of flows. For new flows, the decision is based on traffic descriptor parameters and on measurements of DiffServ class load against a given class threshold (the initial bandwidth allocation for that class). The way the model should be parametrized has already been presented and discussed in a previous paper [13]. In addition, this paper also presents an extension proposal to account for global mobility, where MNs can make handovers across different MAPs and domains.

When speaking of handover flows, the decision is based on inputs from MN's QoS context at pAR and on measurements of MN's DiffServ classes load in nAR at the time of the handover against a given class threshold.

Additionally, and if necessary, the dynamic allocator, which is a bandwidth reallocation mechanism, dynamically redistributes the allocate bandwidth for class with best-effort traffic among classes with stricter QoS requirements in order to accommodate more incoming handover flows in higher priority classes.

The main objective of the dynamic allocator is to increase the resources utilization. This mechanism enhances the admission control process with an adaptive behavior to the network's conditions allowing a more appropriate response to handover flows QoS requirements while increases the resource utilization without compromising the global system quality and stability.

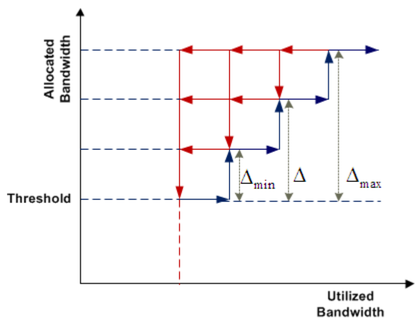


Figure 2: Allocated Class Bandwidth with Hysteresis

Figure 2 illustrates the dynamic allocator's reallocation mechanism behaviour. Making bandwidth reallocations in fixed step sizes the implemented scheme provides a very predictable and stable behavior to the reallocation mechanism (see equation 1).

$$\#steps_i = int\left(\frac{(Class_i + ClassCntxt_i) - T_i}{\triangle min_i}\right) + 1 \quad (1)$$

With the dynamic allocator, the admission control algorithm accepts the MN's handover flows if there is available bandwidth to reallocate in the class ($\triangle max_i$) in question. For instance, assuming that an MN starts with handover procedure to move to a new AR, and at that moment, for MN's class i to be admitted in the new AR, the number of steps that must be reallocated is $\#steps_i = 3$, in such case the reallocated bandwidth must be

$$\triangle Class_i = 3 \times \triangle min_i,$$

but there will only be reallocation if and only if the $3 \times \triangle min_i \leq \triangle max_i$.

For call setup, a two-way signaling protocol is used for new applications to express their service requests to the network. Service requests contain a traffic descriptor describing the worst case application traffic behavior and the required DiffServ class.

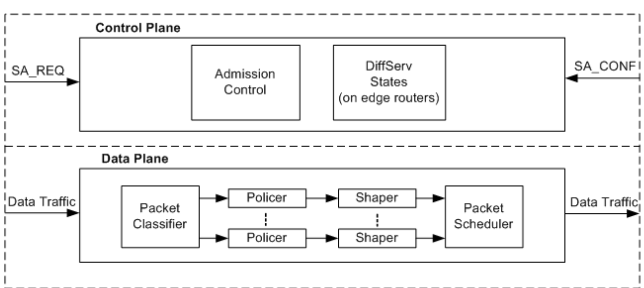


Figure 3: Communication Process with Edge Routers

Signaling protocol communicates with edge router Signaling Agents (SAs) the traffic and service specification of an incoming flow (see Figure 3). To communicate new flows, the Correspondent Node (CN) uses its SA to request services from the network; this SA is responsible for the delivery of all service request messages. Signaling Request (SA-REQ) messages sent by CN contain the traffic description that will be the RMF input. The message contains two parameters: Desired Bandwidth and DiffServ Class. The Signaling Agent sets the desired bandwidth and class such that each SA on path could read and pass those parameters to the resource management function. If one of the edge routers in the path fails to satisfy the desired QoS, the receiving Signaling Agent generates a negative Signaling Confirmation (SA_CONF) message to the SA initiator (the CN) with a negative decision, and the flow is aborted. Otherwise, the receiving Signaling Agent sets the SA_CONF with a positive decision.

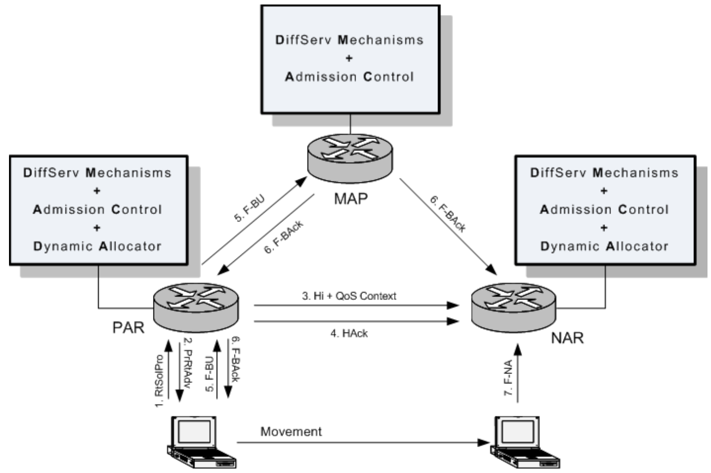


Figure 4: Signaling Procedure for Handover Flows

Figure 4 shows the signaling procedure for intra-domain handovers. For intra-domain handovers, the MN's QoS Context at pAR is conveyed by HI messages to nAR. The HI messages will be handled by the RMF of nAR. The HI handover signaling message triggers the resource management functionalities in the nAR before the handover occurs resulting in a proactive behavior of the RMF, enabling RMF to configure itself automatically for incoming handover flows.

Therefore, (Step 1.) whenever an MN wishes to change its point of attachment, it must ask for a new CoA address to nAR by sending Router-Solicitation-for-Proxy (RtSolPro) message to pAR. (Step 2.) The pARs receives the RtSolPro message and generates a Proxy-Router-Advertisement (PrRtAdv) message with a prospective new MN CoA and sends it to the MN. (Step 3.) The pAR also creates an HI message containing the nAR address as well as the MN's QoS context to be sent to nAR. The MN's QoS context in the pAR is extracted from the rate estimator of the resource management function which measures each DiffServ class bandwidth in use on pAR by MN at that time. This per-Class state information (MN's QoS context) is stored in the mobility options of the HI message´s field. (Step 4.) The new access router receives the HI message and in turn processes the mobility and corresponding resource management functions. The resource management function decides which MN's DiffServ classes of flows it is able to accept and, if necessary, the dynamic allocator of resource management function fetches more bandwidth for classes with stricter QoS requirements to accommodate the flows belonging to those priority classes.

Next, it forms a valid Care-of-Address (CoA) or validates the prospective new CoA and places the CoA and the admission control decision on a HAck message, and returns the message to the pAR. The pAR receives the HAck, validates the new CoA address and sends a negative decision on an SA_CONF message (the message is not illustrated in the Figure, containing the rejected flows to CN. (Step 5.) Then, the MN sends a Fast Binding Update (F-BU), via pAR to MAP for binding its previous CoA to the new CoA. (Step 6.) MAP receives an F-BU message and sends a F-BAck message to MN and nAR. The MN needs to wait for the F-BAck message before it makes the handover because this message indicates that MAP is prepared to make the tunneling of the packets to nAR. When the MN receives the F-BAck message, it first disconnects from the pAR and then re-attaches to nAR. (Step 7.) Once in the nAR, MN sends an FNA message to receive the buffered packets in the nAR and registers its new CoA with HA and CNs by sending a binding update message.

This proactive (before MN moves to a new location) and dynamic (by adjusting the load within classes for handover flows) RMF

PAR (85%)

Class 1	Class 2	Class 3	Class 4	
12	36	30	24	%
42	126	105	84	kbps
5.25	15.75	13.3	10.5	(kbps) $\times$ 8 MN

(a) At pAR

NAR (15%)

Class 1	Class 2	Class 3	Class 4	
6	12	18	18	%
21	42	63	63	kbps
10+11	20+22	30+33	30+33	(kbps)(1 + 1) MN

(b) At nAR

Table I: Generated Traffic

behavior pretends to provide seamless mobility to mobile users running real-time applications by maintaining the same QoS level across ARs.

IV. Model Performance&Results

In this section it will be analyzed the architecture performance when it is subjected to CBR and exponential (EXP) traffic, and also its behavior with Priority queue (PRI) and Weighted Round Robin (WRR) scheduling algorithms which are the most common queueing disciplines used in the DiffServ architecture.

The model has been implemented in the Network Simulator version 2 (ns-2), patched with IEEE 802.21, HMIPv6 and FMIPv6 extensions [16], [9]. The model performance when all traffic is CBR has been already analyzed and discussed in a previous work [14].

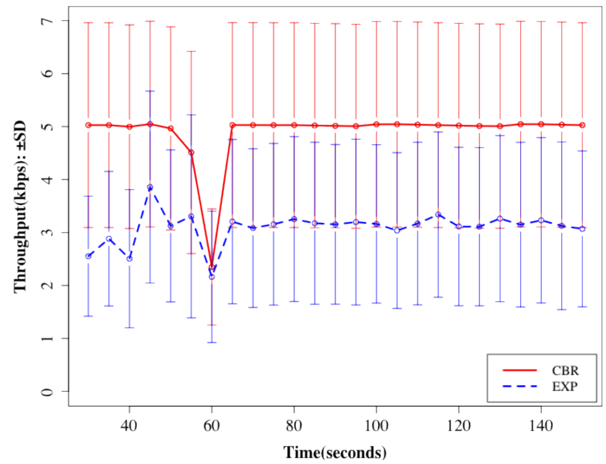
Before starting with the analysis of the results, it is important to highlight the essential characteristics of the PRI and WRR queueing disciplines, so that the results can be better understood.

In packet-switched networks, the packets belonging to various flows are queued for transmission at an output buffer. The manner in which they are selected for transmission on the link is known as queueing scheduling discipline. There exist two common queueing disciplines for the DiffServ architecture which are PRI and WRR. In the PRI queueing discipline the packets arriving at a router are classified and forwarded for its output priority queue. The priority queueing discipline always chooses to transmit first the packet that belongs to the highest priority class.

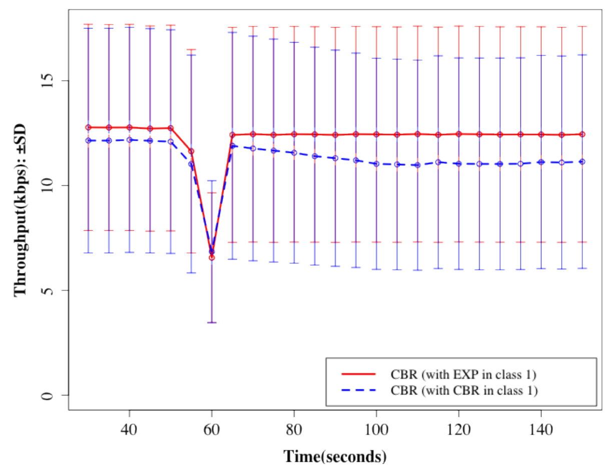
In the WRR queueing discipline, the scheduler alternates the service among the classes in a circular manner serving the highest priority class first. Each class may have a different amount of service which is assigned by means of a weight. This kind of discipline is called "work-conserving" because it never allows the link to be idle if there are packets queued, to be transmitted, in any class.

This work mainly deals with the class of applications which allows people to use audio/video to communicate with each other in real time. In this class, delays smaller than 150 milliseconds are unperceived by a human listener and delays between 150 and 400 milliseconds could be acceptable.

In order to transmit voice over Internet, the analog audio signal must be converted or encoded into a digital signal and then compressed to reduce the bit rate of the stream. There are several compression schemes. The most common are GSM (13 kbps), G.729 (8 kbps) and G.723.3 (both 6.4 and 5.3 kbps). In order to simulate traffic voice, the flows generated for class 1 have been modeled with CBR and EXP traffic transmitting at a rate of 5.3 kbps (for EXP traffic when it is in the period ON), which represent voice traffic encoded with G.723.3. The exponential ON-OFF-traffic model has been set up with the periods ON and OFF of 650ms and 350ms, respectively [6].



(a) CLASS 1



(b) CLASS 3

Figure 5: MN_1 traffic with PRI scheduling

By applying the AD Policies to new flows one gets:

PAR

Class 1	Class 2	Class 3	Class 4	
10	20	30	24	%
31.5	63	93.1	84	kbps

(a) At pAR

NAR

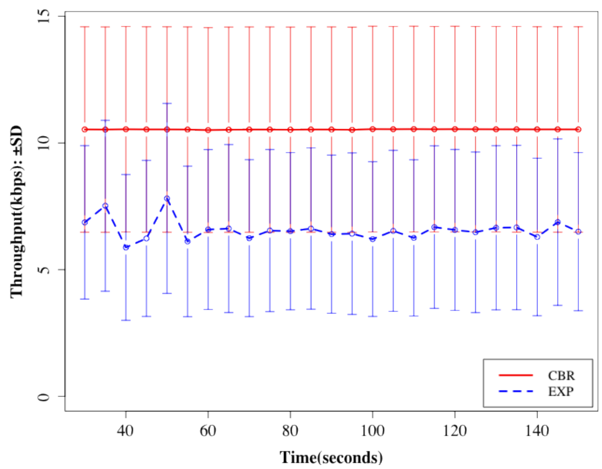
Class 1	Class 2	Class 3	Class 4	
6	12	18	18	%
21	42	63	63	kbps

(b) At nAR

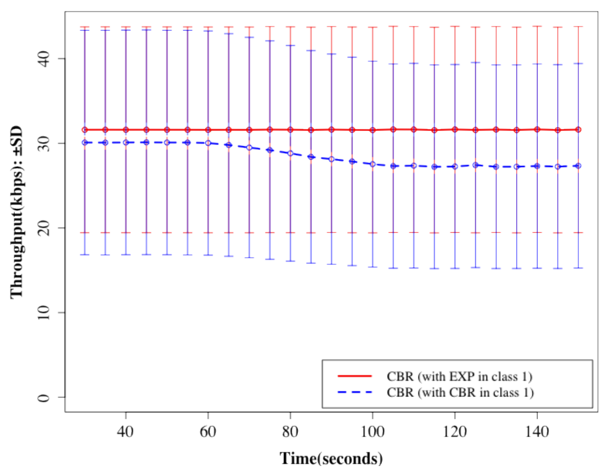
Table II: Admission Control Process for New Flows

Figure 7 shows the simulation scenario that has been used. The simulation scenario consists in a part with the global Internet, where ten CNs and HA's are located, and another part with a F-HMIPv6 aware DiffServ domain with two access routers and ten MNs.

The MNs are receiving data from CNs located at another DiffServ domain of the global Internet. The traffic transmission is one to one (CN→MN) and each CN is generating four flows marked with different DiffServ Code Points (DSCPs). Therefore, forty flows have been generated in the total. Eight of the ten MNs are initially located in pAR and others two MNs are fixed at nAR. MNs are moving



(a) CLASS 1



(b) CLASS 3

Figure 6: MN_10 traffic with PRI scheduling

randomly to nAR in a range time between 50 and 100 seconds, and the network load on nAR after MNs handovers is 131.6%. Having this scenario in consideration, the tests have been conducted in order to evaluate the voice traffic models and the queueing disciplines in the traffic belonging to a mobile node, in movement to nAR, and in a stationary mobile node at nAR. The mobile nodes have been denominated as MN_1 and MN_10, respectively.

After the handover of $8 \times$ MN in pAR to the nAR one gets:

PAR

Class 1	*Class 2*	*Class 3*	*Class 4*	
9	18	26.6	24	%
31.5	63	93.1	84	*kbps*

(a) At pAR

NAR

Class 1	*Class 2*	*Class 3*	*Class 4*	
15	30	44.6	42	%
131.6% of overload				

(b) At nAR

Table III: Admission Control Process for Handover Flows

This scenario has been configured in four different manners: 1) with CBR traffic and PRI queueing discipline; 2) with CBR traffic

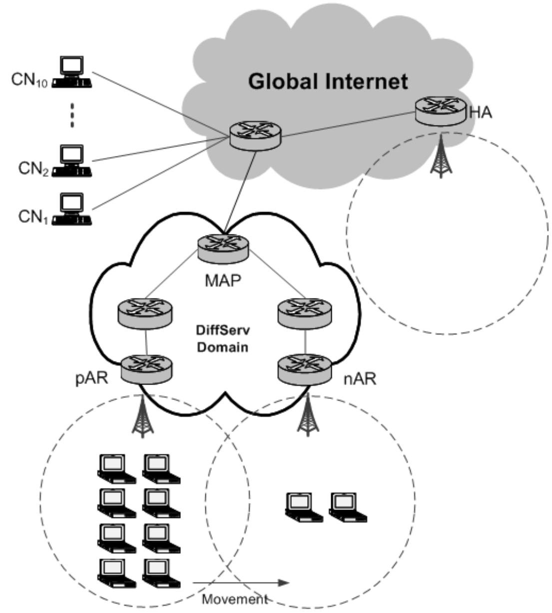


Figure 7: Simulation Scenario

and WRR queueing discipline; 3) with EXP traffic in class 1 and WRR queueing discipline; 4) and with EXP traffic in class 1 and PRI queueing discipline. The traffic generated in the four classes of each mobile node are described in the Tables I, II and III. These tables show the envisioned traffic load, for the traffic generated in the simulations, on pAR and nAR, after applying the admission control to new flows and after MNs have completed the handover. The proposed model is able to work rightly under any traffic load condition, however, the model can be more efficient if the $\triangle min_i$ parameter is tuned to the traffic dynamics.

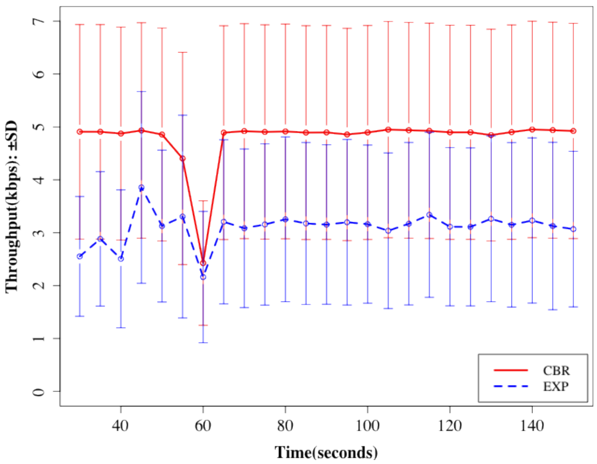
Figures 5 and 6 show the traffic behavior in classes 1 and 3 for MN_1 and MN_10, when the scenario is configured with CBR traffic and PRI queueing discipline, and with EXP traffic and PRI queueing discipline, respectively.

As expected, figures show that the influence of EXP traffic on the transmission rate of the other flows is insignificant and shows only small differences which eventually should be due to the EXP bursty nature which implies additional retransmissions at the MAC level. Nevertheless, the MN_10' traffic (Fig. 6) can maintain its transmission rate in the two higher priority classes even in the presence of EXP traffic, only having slightly decrease in the lowest priority class.

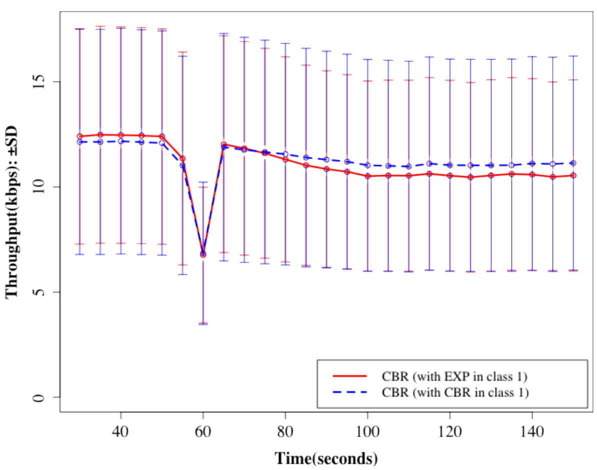
Figures 8 and 9 show the traffic behavior in classes 1 and 3 for MN_1 and MN_10, when the scenario is configured with CBR traffic and WRR queueing discipline, and with EXP traffic and WRR queueing discipline, respectively.

The Figures show that with WRR queueing discipline, the mean rate oscillates a little more for both types of traffic however, with CBR traffic in class 1, the mean rate has been slightly reduced in the higher priority classes while with EXP traffic in class 1, the mean rate for classes 1 and 3 is similar that obtained in the Figs 5 and 6.

As the WRR scheduler alternates the service among the classes in a circular manner, providing each class a certain amount of service, the CBR traffic can only transmit during its assigned amount of service. After this has occurred, it must wait for its turn to transmit again. This causes a slight reduction in the mean rate, mainly for lower priority classes because they have a smaller weight.

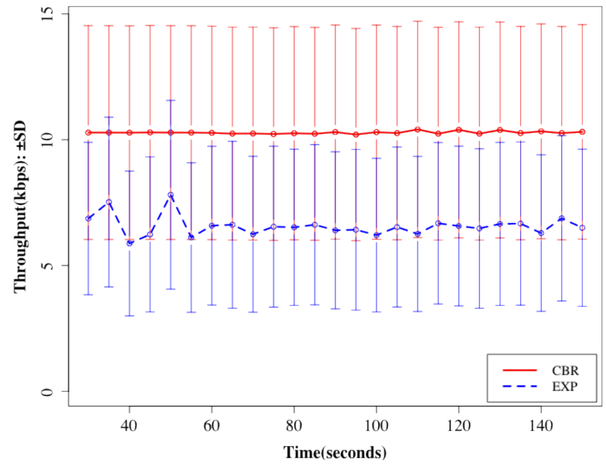


(a) CLASS 1

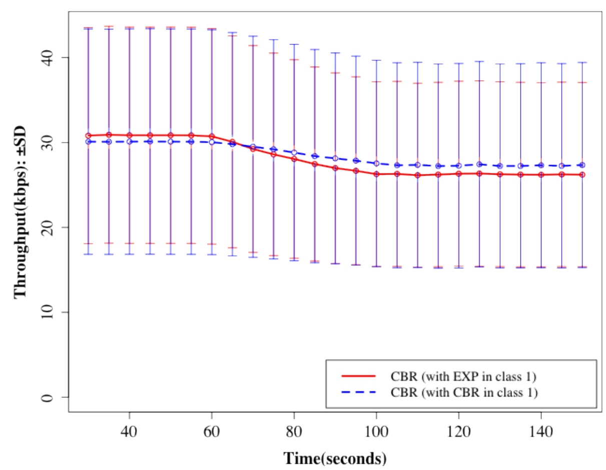


(b) CLASS 3

Figure 8: MN_1 traffic with WRR scheduling



(a) CLASS 1



(b) CLASS 3

Figure 9: MN_10 traffic with WRR scheduling

V. Conclusion

The future architecture of the Internet, in addition to support global mobility, will also enable multimedia applications to explicit their requirements of network service with the quality desired by the respective applications, and consequently the network instead of just providing an unique service, the best service (best- effort), will support multiple service classes that provide probabilistic performance guarantees to multimedia applications.

This paper proposes a solution for improving the resource management system that gives support to the mobility of a mobile device (such as the iPad, PDAs, iPhones and other cell phones) by properly and dynamically managing the available resources, and also endows the network service with multi-class probabilistic quality guarantees to the future multimedia applications.

This new solution for management of mobile resources is well adapted to mobile environments with high dynamics, thus it has adaptive features which are achieved through the exchange of signaling messages and a distributed resource management.

In what respects to simulations results, the results are very similar for both types of traffic, as expected because the QoS mechanisms of the proposed model protect traffic belonging to the priority classes regardless of the traffic type. The small differences between CBR and EXP traffic can eventually derive from the fact of the 802.11 error control, the ARQ error control mechanism, makes additional retransmissions with EXP traffic. The PRI queueing discipline has shown to be more efficient for CBR traffic belonging to a priority class because on the contrary to the WRR queueing discipline, the packet is served according to is priority and do not need to wait for its queue turn to be served when a packet arrives at a priority queue.

References

[1] A R. Koodli, E. *Mobile IPv6 Fast Handovers*, rfc 5568 ed. IETF, July 2009.

[2] Awduche, D., and Agu, E. Mobile extensions to RSVP. In *Proceedings of the Sixth International Conference on Computer Communications and Networks* (1997), pp. 132–136.

[3] Awduche, D. O., and Agu, E. Mobile extensions to RSVP. In *IC3N '97: Proceedings of the Sixth International Conference on Computer Communications and Networks* (Washington, DC, USA, 1997), IEEE Computer Society, p. 132.

[4] Belhoul, A., Sekercioglu, Y. A., and Mani, N. Mobility-Aware RSVP: A framework for improving the performance of multimedia services over wireless IP-based mobile networks. *Computer Communications 32*, 4 (March 2009), 569–582.

[5] Belhoul, A., Y., A., and Mani, N. Mobility protocols and RSVP performance in wireless IPv6 networks: shortcomings and solutions. *Wireless Communications and Mobile Computing 8*, 9 (2008), 1183–1199.

[6] Dang, T. D., Sonkoly, B., and Molnar, S. Fractal analysis and modeling of VoIP traffic. In *Telecommunications Network Strategy and Planning Symposium. NETWORKS 2004, 11th International* (june 2004), pp. 123 – 130.

[7] FANKHAUSER, G., HADJIEFTHYMIADES, S., NIKAEIN, N., AND STACEY, L. Interworking RSVP and Mobile IP Version 6 in Wireless Environments. *Proc. 4th ACTS Mobile Communications Summit* (1999).

[8] FU, X., KARL, H., AND KAPPLER, C. QoS-Conditionalized Handoff for Mobile IPv6. In *NETWORKING 2002: Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications*, E. Gregori, M. Conti, A. Campbell, G. Omidyar, and M. Zukerman, Eds., vol. 2345 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2006, pp. 721–730.

[9] HSIEH, R., AND SENEVIRATNE, A. A comparison of mechanisms for improving mobile IP handoff latency for end-to-end TCP. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking* (New York, NY, USA, 2003), ACM, pp. 29–41.

[10] HUANG, N.-F., AND CHEN, W.-E. RSVP extensions for real-time services in hierarchical mobile IPv6. *Mob. Netw. Appl. 8*, 6 (2003), 625–634.

[11] LEU, S.-J., AND CHANG, R.-S. Integrated service mobile internet: RSVP over mobile IPv4/6. *Mob. Netw. Appl. 8*, 6 (2003), 635–642.

[12] LIU, C., LIU, Y., QIAN, D., AND LI, M. An Approach of End-to-End DiffServ/MPLS QoS Context Transfer in HMIPv6 Net. In *Autonomous Decentralized Systems, 2007. ISADS '07. Eighth International Symposium on* (march 2007), pp. 245 –254.

[13] LOPES, N. V., NICOLAU, M. J., AND SANTOS, A. A Micro-Mobility Solution for Supporting QoS in Global Mobility. *18th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2010)* (23-25 September 2010). (Best Student Paper Award).

[14] LOPES, N. V., NICOLAU, M. J., AND SANTOS, A. A QoS-Enabled Resource Management Scheme for F-HMIPv6 Micro Mobility Approach. *Journal of Telecommunications Systems* (2011). Accepted in April 2011.

[15] MAMELI, R., AND SALSANO, S. Use of COPS for Intserv operations over Diffserv: architectural issues, protocol design and test-bed implementation. In *IEEE International Conference on Communications, 2001 (ICC 2001).* (2001), vol. 10, pp. 3265–3270.

[16] NIST. *The Network Simulator NS-2 NIST add-on (IEEE 802.21 model)*. National Institute of Standards and Technology, January 2007. http://w3.antd.nist.gov/seamlessandsecure/ (last visited 05/2011).

[17] SOHAIL, S., AND JHA, S. The Survey of Bandwidth Broker. Technical Report UNSQ-CSE-TR-0206, University of New South Wales, Sydney, Australia, May 2002.

[18] SOLIMAN, H., CASTELLUCCIA, C., ELMALKI, K., AND BELLIER, L. Hierarchical Mobile IPv6 (HMIPv6) Mobility Management. RFC 5380 (Proposed Standard), 2008.

[19] TSENG, C.-C., LEE, G.-C., LIU, R.-S., AND WANG, T.-P. HMRSVP: A Hierarchical Mobile RSVP Protocol. *Wireless Networks 9* (March 2003), 95–102.

[20] WAN, Z., PAN, X., AND PING, L. QoS Provisioning in an Enhanced FMIPv6 Architecture. In *International Conference on Computational Science and its Applications (ICCSA 2005)* (May 2005), vol. 3481, pp. 704–713.

[21] ZHANG, R., AND LONG, K. QoS issues in mobile IP: challenges, requirements and solutions. In *ICCC '02: Proceedings of the 15th international conference on Computer communication* (Washington, DC, USA, 2002), pp. 802–812.