Journal of Advances in Science and Technology
Vol. 15, Issue No. 1, March-2018, ISSN 2230-9659

Resource Management in PMIPv6 Networks

Dr. Sridevi*

Department of Computer Science, Karnatak University, Dharwad

Abstract – An aim of the proposed solution is to address the issue of resource management in PMIPv6 networks which is divided into two parts. First, a class based admission control is proposed where real-time traffic are given more priority by being assigned more resource units and non-real-time traffic are allowed to occupy fewer resource units. In the second part of the solution, a modified version of the routing approach proposed in [9] is implemented on top of the admission control scheme and the performance of the joint routing solution and proposed admission control is investigated.

Keywords: PMIPv6, MN, LMA, Mobility, etc…

\- X -

1. INTRODUCTION

Rapid growth in number of mobile devices such as smart phones, PDAs, and laptop computers demand for anytime anywhere" high-speed Internet access. All-IP mobile networks are a combination of Internet and telecommunication networks which provide better interworking between heterogeneous radio access networks and support ubiquitous communication [2]. Next generation of mobile networks are expected to support a variety of mobile devices over IP based mobile networks. As a result the need to have high-speed Internet access and an intelligent mobility management support along with QoS mechanisms is ever increasing.

PMIPv6 have several advantages over MN based micro mobility protocols such as Cellular IP [3], HMIPv6 [5], FMIPv6 [4] [6] and macro mobility solutions such as MIPv6 [1]. Some of the benefits include minimizing the handover latencies, reducing the overheads such as signaling over the wireless link and non-complex deployment given that there is no need to upgrade the protocol stack of MNs [7]. Several drawbacks are associated with triangular routing from the CN to the LMA and then to the MN; such as routing issues, tunneling overheads, and creating congestion around the LMA. These drawbacks can immensely affect the network performance, for instance the bottleneck effect caused by the congestion around the LMA reduces the capacity of the access network and the overall throughput of the network. Thus, managing the resources at the LMA becomes highly important.

In order to address the resource management issues at the LMA,a class-based resource allocation scheme is considered. Real-time traffic is given more priority by being able to simultaneously occupy more resource units, whereas non-real time traffic can use less resource units. As mentioned earlier, the tunneling based nature of PMIPv6 can cause routing issues, reduce the overall capacity of the network and result in the LMA/s becoming the main bottle neck node/s in the network. Allowing all types of traffic to be routed through the LMA uses the valuable available bandwidth at this node, especially in case of routing non-real time IP traffic, such as Peer to Peer (P2P) downloads, through the LMA. Under this perspective, a cut-down version that does not consider the effect of mobility, the QoS aware Dynamic Route Optimization (DRO) and routing policy proposed in [49] is implemented, where sessions that are not delay sensitive are to bypass the LMA and establish a direct binding with the CN.

The following issues were addressed through investigating two sets of scenarios:

1. Categorizing traffic into two classes: Real-time or non-delay tolerant, and Non real-time or delay tolerant. The PMIPv6 network is modeled as a tandem queuing network, where each node operates as an M/M/m/m queue or Erlang Loss system [8] with two types of arrival process. The capacity of each node is assumed to be divided into m resource units.

2. The proposed admission control scheme differentiates between each class of traffic by means of allocating different number of resource units; each real-time session will benefit from the simultaneous use of several resource units whilst each non-real-time session can occupy one of the m resource units at a time.

3. The state transition diagram of one node is analyzed.

4. In the first scenario, sessions from both classes of traffic traverse through the LMA node which is in tandem with the MAGs.

5. In the second scenario, a cut-down version of the RO and policy proposed in [9] is implemented, where real-time traffic are routed through the LMA and non-real-time traffic are routed directly to the MAGs.

2. PROBLEM STATEMENT

PMIPv6 provides mobility to the IP devices without their involvement, and it can achieve this by relocating the mobility management functions from the MN to the network. Traffic originated from or sent to an MN is routed through tunnels between the MN's corresponding MAG and LMA or vice versa. This results in high band width consumption and bottleneck effect as well as increased processing time at the LMA. Considering this and the huge expected rise in mobile IP traffic in future, the need for employing an efficient resource management scheme in PMIPv6networks is increasing. Two approaches can be used to address the issue of resource management in PMIPv6 networks:

1. An effective admission control scheme

2. An optimal routing solution..

3. SYSTEM MODEL

PMIPv6 network is modeled as a tandem queuing network. Capacity of each node is assumed to be divided into m resource units which is equivalent to m servers and each node operates as an independent M/M/m/m queue or Erlang Loss system [8]. The primary assumption in future all-IP networks is that n classes of traffic are present, but as mentioned earlier aggregating ows into a few number of classes according to their QoS requirements is much simpler traffic management task than providing QoS for each of the n individual ows. Moreover, with majority of today's Internet traffic being background and being less sensitive to delay (such as downloads of games, music, videos and etc) [10][11], it is wise to have a more general view when it comes to differentiating the classes of traffic in the network. In this respect, two classes of traffic are considered as follows:

I) Delay tolerant or non-real-time traffic and

II) Non delay tolerant or real-time traffic. A similar distinction between different classes of traffic was made in [9]and [12].

Sessions from both classes of traffic have to traverse through the LMA node and if both classes of traffic are treated the same, it may result in QoS disruption.

The proposed admission control scheme, takes a class-based approach when it comes to resource allocation and operates as follows: assuming that the bandwidth required by the real-time traffic is much larger than the bandwidth required by the non-real-time traffic, each real-time session will benefit from the simultaneous use of several resource units whilst each non real-time session can occupy one of the m resource units at a time.

The amount of bandwidth required by the real-time and non-real-time traffic is a and b Kbps respectively. Assume that the capacity of all nodes in the network is equal to each other and is denoted by CT. Assuming that a >> b, then it can be stated that the total bandwidth in each node is equal to m servers or resource units, m being equal to η as shown in equation 2. In the proposed model, admitting a nonreal-time traffic requires one of the resource units and admitting a real-time traffic requires the simultaneous use of several resource units. Number of required resource units by the non-real and real-time traffic are shown by *mb* and *ma* respectively, where *mb* = 1 and *ma* = a=b. Considering one class of traffic at a time and assuming that there is no arrival from the other class, it can be stated that the number of real-time traffic that can be admitted at each node is equal to and the number of non-real-time traffic that can be admitted at each node is equal such that:

$$\Upsilon = C_T/a \qquad (1)$$

$$\eta = C_T/b \qquad (2)$$

One of the primary assumptions in this paper is that there is an identical service time offered in each node of the network with higher service time for the real-time traffic, this assumption restores the independence of inter-arrival times and packet lengths; hence the Klein rock independence approximation is valid and nodes in the network can be modeled as independent M/M/m/m queuing systems [8].

Another valid assumption is the offline route discovery; i.e. admitting the sessions at the gateway is performed with the prior knowledge of the bottlenecks in the network. Dijkstra algorithm is used to find the set of shortest paths in the networks.

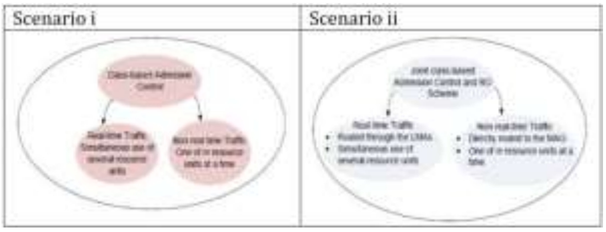
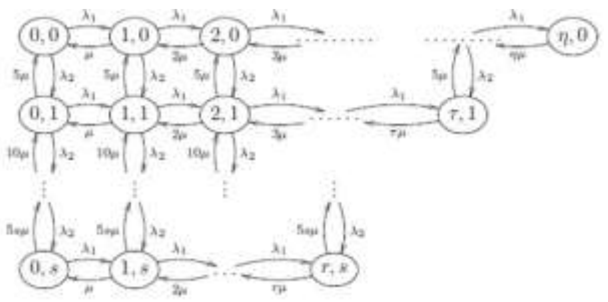


Figure 1: Overview of the proposed resource management scheme for PMIPv6

4. STATE TRANSITION DIAGRAM

State transition diagram for one node in the first scenario is analyzed. State (i; j) means that i requests of non-real-time and j requests of real-time class are present in the system. Let λ_1, λ_2 and $\mu, 5\mu$; be the average call arrival and mean service time of non-real-time and real-time traffic respectively. Consider an LMA node within a PMIPv6 network with capacity C_T , this can be analyzed by a 2-dimensional Markovchain [13] as follows:



Let P(i; j) be the steady-state probability of system being in state (i; j). The probability of system being in an equilibrium state can be found by solving the general balance equation (flow in= flow out) for an internal state:

$$\lambda_1 p(i-1,j) + \lambda_2 p(i,j-1) + (i+1)\mu p(i+1,j) + (j+1)5\mu p(i,j+1)$$
$$= j5\mu p(i,j) + i\mu p(i,j) + \lambda_2 p(i,j) + \lambda_1 p(i,j) \qquad (3)$$

and by the method suggested in [13], this is shown below :

$$p(i,j) = \frac{\frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j}{\sum_{j=0}^{\eta} \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j \sum_{i=0}^{\eta - j m_a} \frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i} \qquad (4)$$

In the state diagram, horizontal arrows to the right and left correspond to arrival and departure of non-real-time traffic into the system. It can be concluded that anon real-time traffic is blocked if all the η servers are busy and this occurs when the system is at the rightmost of any row. Using equation 4 and summing the probabilities of the s + 1 states gives the blocking probability of this class of traffic, where $P_{bnon\text{-}rt}$ denotes the probability of a non-real-time arrival being blocked or denied access to the LMA:

$$P_{b_{non-rt}} = p(0,0)\left[\frac{1}{\eta!}\left(\frac{\lambda_1}{\mu}\right)^{\eta} + \frac{1}{(\eta - m_a)!}\left(\frac{\lambda_1}{\mu}\right)^{\eta - m_a}\left(\frac{\lambda_2}{5\mu}\right) + \cdots + \frac{1}{(\eta - s m_a)!}\left(\frac{\lambda_1}{\mu}\right)^{\eta - s m_a}\frac{1}{s!}\left(\frac{\lambda_2}{5\mu}\right)^{s}\right]$$

$$P_{b_{non-rt}} = \frac{\sum_{j=0}^{s} \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j \frac{1}{(\eta - j m_a)!}\left(\frac{\lambda_1}{\mu}\right)^{\eta - j m_a}}{\sum_{j=0}^{s} \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j \sum_{i=0}^{\eta - j m_a} \frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i} \qquad (5)$$

Vertical arrows up and down, represent arrival and departure of real-time traffic into the node. A real-time arrival is blocked if at least Y+ 1 servers are busy, where Y=η-m_a, in other words when less than m_a resource units are idle or available. In order to calculate the blocking probability of real-time traffic $P_{b_{rt}}, \eta + 1$ probabilities are summed as follows:

$$P_{b_{rt}} = \sum_{i=\eta - m_a + 1}^{\eta} p(i,0) + \sum_{i=\eta - 2m_a + 1}^{\eta - m_a} p(i,1) + \cdots + \sum_{i=\eta - s m_a + 1}^{\eta - (s-1) m_a} p(i,s-1) + \sum_{i=0}^{r} p(i,s)$$

The probability that traffic of type real-time is blocked from the system is:

$$P_{b_{rt}} = \frac{\sum_{j=0}^{s-1} \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j \sum_{i=\eta-(j+1)m_a+1}^{\eta - j m_a} \frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i}{\sum_{j=0}^{s} \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j \sum_{i=0}^{\eta - j m_a} \frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i} + \frac{\frac{1}{s!}\left(\frac{\lambda_2}{5\mu}\right)^s \sum_{i=0}^{r} \frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i}{\sum_{j=0}^{s} \frac{1}{j!}\left(\frac{\lambda_2}{5\mu}\right)^j \sum_{i=0}^{\eta - j m_a} \frac{1}{i!}\left(\frac{\lambda_1}{\mu}\right)^i} \qquad (6)$$

5. SIMULATION SCENARIOS

First Scenario: All through the LMA

In this scenario, sessions from both classes of traffic have to traverse through the LMA node. This means that the set of paths that contain the LMA node are used and set of $\bar{P}$ paths are left underutilized. Nodes within the network are modeled as z independent two-dimensional M/M/m/m queues, this is illustrated in figure 2. Each node can be analyzed by the same state transition diagram.

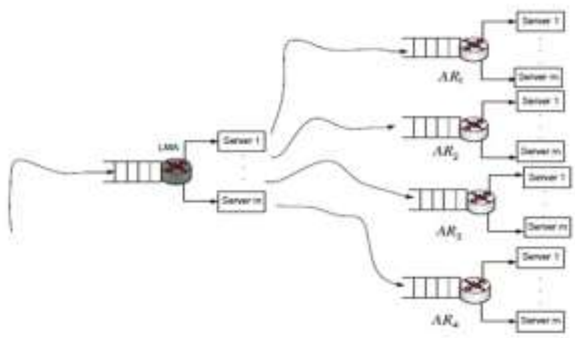


Figure 2: Scenario I: zindependent M/M/m/m queues

Looking at the system from the downlink and considering Figure 2, it can be concluded that the only bottleneck in the system is the LMA node and each session has to pass through two independent queues in tandem before it reaches its destination. This means that a session travels one of these four routes: LMA-AR1, LMA-AR2, LMA-AR3, or LMA-AR4. The admission control scheme for the downlink works as explained below, the same principle stands for the reverse direction (uplink):

Second Scenario: Using RO schemes

An RO scheme similar to the QoS aware DRO, is implemented that operates alongside the proposed admission control scheme from the first scenario. Session's mobility rate is not taken into consideration, and the routing policy works as follows: Upon arrival of a ow at the gateway, based on the ow's class of the traffic, a decision as to route the ow directly to the specific MAG or the LMA/s is made. If the session is non-real-time and it is decided to be routed directly to the MAG, the least cost path from the _ P set of paths is calculated and the non-real-time ow is routed through the least cost path. Calculating the least cost path for a real-time session is done in two stages:

1. The least cost path from the gateway to the LMA k is calculated and the ow is routed along this path till it reaches LMA k.

2. The least cost path from the LMA k to the specific MAG is calculated and the ow is routed through this path.

Looking at the joint admission control and RO mechanism from the downlink direction, it can be stated that though sessions from each class of traffic are routed through two separate paths, they still share resource units at the MAGs. Therefore, entry nodes to the network as well as nodes along the path toward the MAGs, can be modeled by independent one-dimensional M/M/m/m queues and MAGs can be modeled as independent two-dimensional M/M/m/m queues. It must be highlighted that at all the nodes in the network the value of mb and ma is the same as it was in scenario one.

6. PERFORMANCE EVALUATION

The proposed solutions are further investigated by means of MATLAB simulations. The main focus is on the results of simulations carried out in the MATLAB but results of the mathematical analysis are presented to verify the simulation environment built for one node.

Blocking Probability Scenario I

To investigate the impact of the proposed admission control scheme on the total and per class blocking probability, the topology is considered. There are one gateway node, four AR nodes (MAGs) and negative intermediate nodes including the LMA node that provide the backhaul routing in the topology. The capacity of a node is denoted by CT and each node is assumed to have equal capacity CT = 20 Mbps.

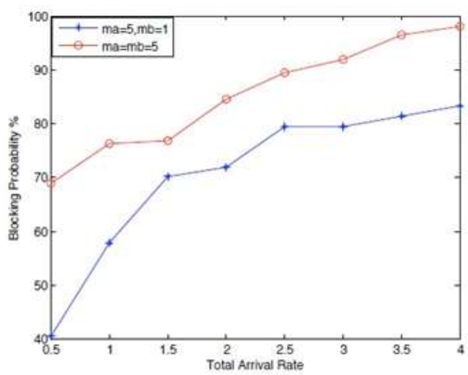


Figure 3: total blocking probability vs arrival rate, where arrival rate is total arrival rate of real time and non-real-time traffic, with 70% of the total arrival rate being non real time at all points

First, the total blocking probability for each arrival rate is computed. The results are then compared to the method where there is no distinction between real-time and non-real-time traffic in terms of number of resource units used i.e. mb = ma = 5, and sessions from both classes of traffic require the simultaneous use of five resource units at a time. As displayed in Figure 3, it can be gathered that the proposed admission control scheme results in a significant decrease in total blocking probability comparing to the case where sessions from both classes of traffic are treated the same.

Next, blocking probability per class of traffic using the proposed admission control is attained. The results are then compared with the scenario where mb = ma = 5.Results in Figure 4 demonstrate that by using the proposed admission control, blocking probability of non-real-time traffic has lowered.

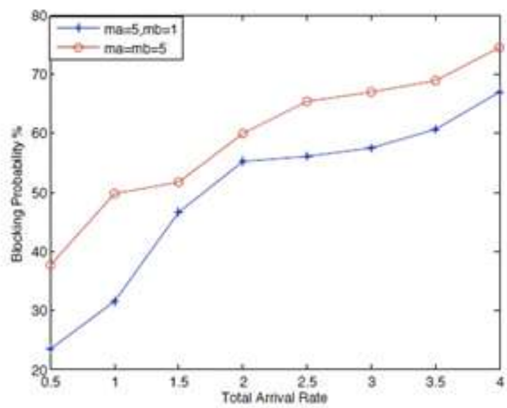


Figure 4: Block probability of non-real-time traffic vs arrival rate, where arrival rate is total arrival rate of real time and non-real-time traffic, with 70% of the total arrival rate being non real-time at all the points

The blocking probability of real-time class of traffic is computed, results of which are compared to the case where mb = ma = 5 in Figure 5. The proposed admission control scheme results in reduction of blocking probability of real time traffic, especially under high congestion in the network. As the total session arrival rate increases, the number of arrivals from the non-real-time class of traffic increases 70% more than the real-time class of traffic. As a result, P_{bnrt} increases and P_{brt} decreases as the total arrival rate increases. Overall it can be concluded that by treating the two classes of traffic differently a great reduction in both total and per class blocking probabilities is achieved.

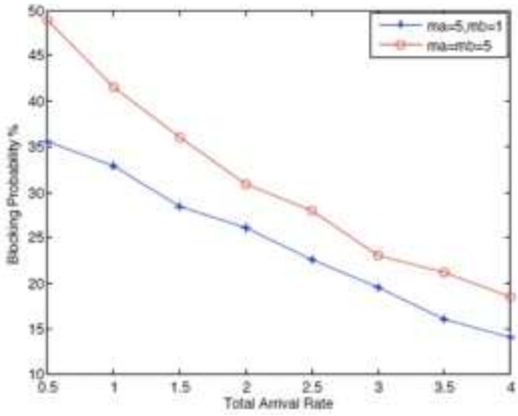


Figure 5: Blocking probability of real-time traffic vs arrival rate, where arrival rate is total arrival rate of real-time and non-real-time traffic, with 70% of the total arrival rate being non real-time at all the points

Blocking Probability Scenario II

The joint performance of the proposed admission control and the modified version of QoS aware DRO is investigated. The objective of employing the QoS aware DRO in conjunction with the proposed admission control scheme is to lower the blocking probability by distributing the traffic through the underutilized paths in the network

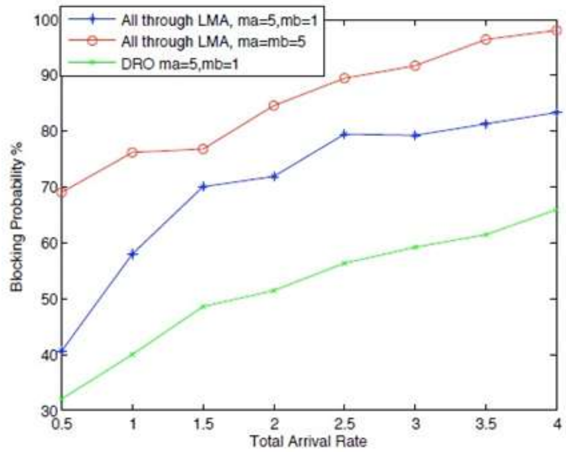


Figure 6: Total blocking Probability vs arrival rate, where arrival rate is total arrival rate of real-time and non-real-time traffic, with 70% of the total arrival rate being non real-time at all the points.

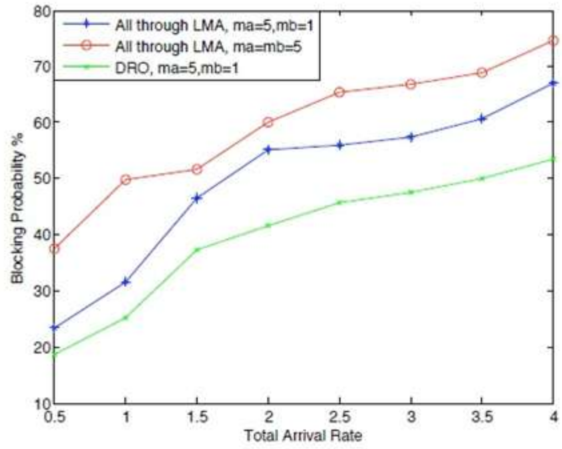


Figure 7: Blocking probability of non-real time traffic vs arrival rate, where arrival rate is total arrival rate of real-time and non-real-time traffic, with 70% of the total arrival rate being non real-time at all points

Using DRO, the non-real-time traffic which forms the majority of the traffic is routed away from the LMA and routed through the underutilized paths. The total blocking probability is computed, results of which are compared to two sets of results previously obtained: the results of scenario one and to the case where all the traffic is routed through the LMA and no distinctions between real-time and non-real-time traffic in terms of resource units allocation is considered (i.e. $m_b=m_a = 5$). As shown in Figure 6, it can be stated that incorporating DRO in conjunction with the proposed admission control results in a significant degree of improvement in terms of reducing the total blocking probability.

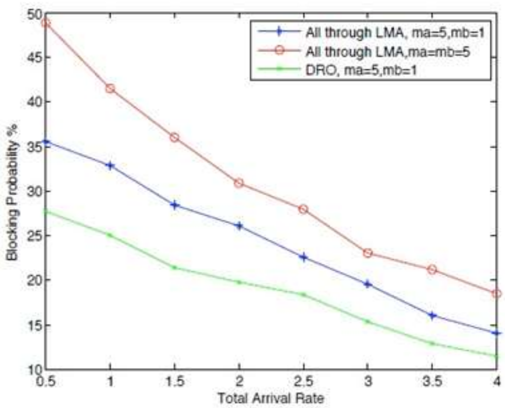


Figure 8: Blocking probability of real time traffic vs arrival rate, where arrival rate is total arrival rate of real-time and non-real-time traffic, with 70% of the total arrival rate being non real-time at all points

Incorporating DRO means that only the real-time traffic has the privilege of using the LMA node and this results in lowering the congestion level at the LMA and reducing the probability of the LMA becoming the main bottleneck node in the network.

Figure 8 illustrates the blocking probability of real-time traffic using three different approaches on the same plot. Real-time blocking probability is at its lowest when using the joint approach. It can be stated that using the joint approach distributes the traffic throughout the network and increases the overall throughput of the network, hence lower total and per class blocking probability is achieved through this approach.

7. CONCLUSION

An issue of resource management in PMIPv6 is addressed with the intention of improving the network performance. First, a class-based admission control scheme is presented to minimize the bottleneck effect caused by triangular routing in PMIPv6. Resource units are rationed amongst different classes of traffic according to their QoS requirements and an analytical model is presented. PMIPv6 network is modeled as an M/M/m/m tandem queuing network with two types (classes) of arrival process: real time and non-real time traffic. The proposed admission control scheme priorities the real time traffic by allowing this type of traffic to use several resource units at a time. Secondly, in order to address the routing inefficiencies in PMIPv6, a QoS aware DRO is implemented which works in conjunction with the proposed admission control. Non real-time traffic that do not require mobility support mechanisms are routed away from the LMA and are routed through the underutilized paths.

REFERENCES

[1] D. Johnson, C. Perkins, and J. Arkko, “Mobility support in ipv6,” RFC 3775, June 2004. 32, 33, 81

[2] I. Akyildiz, J. Xie, and S. Mohanty, “A survey of mobility management in next-generation all-ip-based wireless systems,” IEEE,Wireless Communications, vol. 11, pp. 16 – 28, aug. 2004. 33, 35, 80

[3] A. Campbell, J. Gomez, S. Kim, A. Valko, W. Chieh-Yih, and Z. Turanyi, “Design, implementation, and evaluation of cellular ip,” Personal Communications, IEEE, vol. 7, pp. 42 –49, 2000. 35, 81

[4] R. Koodli, “Fast handovers for mobile ipv6,” RFC 4068, July 2005. 35, 81

[5] H. Soliman, C. Castelluccia, K. El Malki, and L. Bellier, “Hierarchical mobile ipv6 mobility management (hmipv6),” RFC 4140, August 2005. 35, 81

[7] K.-S. Kong, W. Lee, Y.-H. Han, M.-K. Shin, and H. You, “Mobility management for all-ip mobile networks: mobile ipv6 vs. proxy mobile ipv6,” Wireless Communications, IEEE, vol. 15 Issue:2, pp. 36 – 45, April 2008. 37, 81, 84

[8] D. P. Bertsekas and R. Gallager, Data networks (2. ed.). Prentice Hall, 1992. 47, 48, 66, 82, 90, 91, 104

[9] A. D. Pragad, V. Friderikos, P. Pangalos, and A. Aghvami, “Qos aware dynamic route optimization for proxy mobile ipv6 networks,” Wireless Communications And Mobile Computing, vol. 11, p. 508521, 2009. 51, 82, 83, 84, 85, 87, 89, 90, 93, 116, 117

[10] Managing Peer-To-Peer Traffic With Cisco Service Control Technology, Cisco White Paper, 2005. 90, 103 [81] Cisco Visual Networking Index: Forecast and Methodology, 20112016, Cisco White Paper, 2012. 90, 103

[12] A. Pragad, P. Pangalos, V. Friderikos, and A. Aghvami, “Dynamic qos aware route optimization for networks with mobility agents,” in Consumer Communications and Networking Conference, CCNC 2009, 2009. 90 [83] L. Kleinrock and R. Gail, Queueing Systems: Problems and Solutions. John Wiley and Sons, 1996. 95, 96

Corresponding Author

Dr. Sridevi*

Department of Computer Science, Karnatak University, Dharwad