A QoS-Aware Handoff using RSVP in Next Generation Wireless Networks

Buyurman Baykal Özgür B. Akan
Department of Electrical & Electronics Engineering
Middle East Technical University, 06530 Ankara / Turkey
{buyurman,e121467}@metu.edu.tr

Abstract

The handoff call blocking probability minimization is an important issue for seamless wireless network operation. Handoff delay minimization is in parallel with this task. Mobile stations with QoS requirements should be able to receive the same service level during and after handoff. This constraint amplifies the handoff latency caused forced termination problem. This research introduces a handoff method, which performs some preparation phase tasks in advance, in wireless networks. With the Expected Visitor List (EVL) deployment at base stations and RSVP, QoS requirements of an incoming mobile station is obtained and examined, and hence overall handoff latency minimization is achieved. A virtual QoS provisioning method for call admission control of the mobile is also proposed. Performance evaluation with simulation experiments shows that the proposal can be a promising methodology for QoS provisioning in the Next Generation (NG) wireless networks.

1 Introduction

Next Generation (NG) wireless networks should be able to provide users with reliable QoS levels. Handoffs between coverage areas of the same architecture as well as different architectures will require intelligent handoff management to preserve the service level of the mobile stations.

The QoS requirements cannot be satisfied or guaranteed only by local resource availability information in the wireless domain, i.e., base stations, or switching centers. Application specific service requirements should be satisfied in an end-to-end fashion. This necessitates the integration of QoS provisioning and/or resource reservation in wireless and wired parts of the end-to-end traffic flow during handoff.

The preparation latency of the QoS requirements along the new path, however, brings the forced termination risk of active connections during handoff. After the detection of the handoff necessity, next step to go over by the network is to make some preparations. These include performing an admission control algorithm against the owner of the incoming handoff request, and hence (in case of acceptance) performing network preparations for the handoff including route reconstruction for transferred connections, determination of crossover switches and applying a channel assignment strategy for the mobile in its new cell.

With the NG wireless networks, additional preparation phase task is to make end-to-end QoS provisioning and resource reservation (if possible) for the incoming mobile. In current literature, there are studies proposed to provide seamless QoS levels during handoff in wireless networks [1], [2], [3], [4], [5]. All these proposals either requires mobility pattern information, which may not be accurate in heterogeneous next generation wireless networks, or makes some actual reservations in advance, which may result in waste of resources. Resource ReSerVation Protocol [6], is a powerful protocol for setting up reservations for certain QoS specifications along a path. Hence, it can be easily adopted by wireless networks to provide end-to-end QoS guarantees. However, plain implementation of RSVP in wireless networks will bring high overhead in wireless resources during handoff. With the help of an agent deployment at access points of wireless networks, the overhead can be reduced and the strength of RSVP can be benefited.

In this research, a new handoff method is proposed to provide QoS continuity during handoffs in next generation wireless networks. This is tried to be achieved by informing candidate base stations about the QoS requirements of the prospective mobile stations well before the handoff time. Expected Visitor List (EVL) processors at candidate BSs, which are aware of the mobile's QoS expectations, can then execute RSVP-based virtual QoS provisioning procedure to capture the end-to-end resource availability status. The result of the provision query will be stored in a record called EVL entry at the EVL processor. Any change in the resource availability will be reflected to the provision result and the entry will be kept updated to be ready in case of actual handoff. Hence, the incoming mobile will take the advantage of the method if its entry contains up-to-date in-

Proceedings of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)
1530-1346/03 $17.00 © 2003 IEEE

formation. The performance simulations shows that the proposed method is a promising methodology for QoS maintenance in next generation wireless networks.

2 Expected Visitor List (EVL)

2.1 Decomposition of Handoff Messaging

The target cell accepts a handoff request if its resources can support the resource demands of the mobile with certain characteristics at the time of handoff. Otherwise handoff request is simply rejected and the ongoing call is forced terminated. This approach brings about considerable latency since decision making process about the admission of the mobile requires some messaging between the mobile, current base station and the target base station.

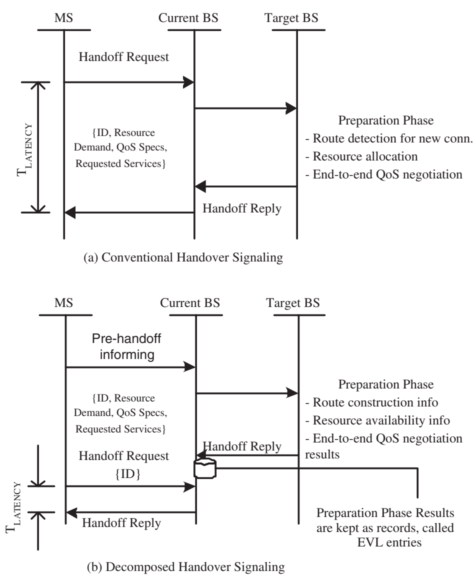

Figure 1. Handoff messaging in existing and new scenarios.

On the other hand, to make some QoS provisioning in advance, QoS expectations of the mobile station must be learned a priori. In order to make the candidate cells aware of expected visitor mobiles in advance, decomposition of handoff messaging is preferable in a way that messaging at the time of handoff as in Fig. 1(a) could be so shortened that it could contain only the identity information of mobile, current and target base stations excluding the pre-transferred information about the mobile demand details as shown in Fig. 1(b). Based on the information received from the decomposed handoff messaging, candidate base stations become capable of initiating the preparations for the incoming mobile in very advance. As a result, much shorter handoff latency than the conventional methods might be achieved.

The classical handoff messaging is decomposed into two parts, one being the non real-time component containing the information about mobile id, traffic specifications, QoS expectations, resource demands making it possible to perform preparations before a possible handoff request. The other part is the real time component and it is the message containing only the mobile id making it possible to execute the handoff based on the preparations.

2.2 EVL Entry Structure

In the proposed method, EVL processors at base stations keep records per neighbor,i.e., EVL entries, to store the decisions and preparations of each expected mobile station. The EVL entries are created first at the EVL processors deployed at the current cell of the mobile. This is because the QoS expectations are captured by the EVL processor at the current base station. A sample EVL entry, as shown in Fig. 2, will include fields for Resource Demand, QoS Expectations, Traffic Specifications, Preemption Priority, and Corresponding Nodes of the mobile station, and the results of the QoS provisioning and connection admission control tasks.

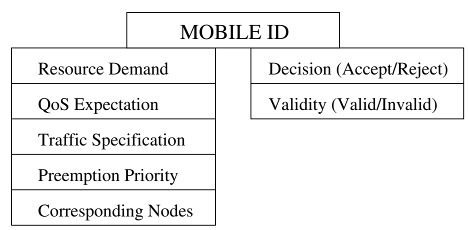

Figure 2. A typical EVL entry structure.

The basic wireless networking illustration with EVL deployed at base stations is seen in Fig. 3. In this example, the EVL_i has entries for all mobiles connected BS_i firstly. The resource demand, QoS expectations, traffic specifications, preemption priority and corresponding nodes data are entered in the appropriate fields of the EVL entry. Since the EVL processor does not process the entries which are current residents of the same cell with itself, the result fields are not entered. Upon the creation of the entries at the original EVL, they are copied to the other EVL processors in the shadow cluster [7], i.e. the cluster of possible target cells, of the mobile station.

The cloning of the EVL entries in the neighbor EVL processors requires signaling between the original EVL processor and the ones in shadow cluster. First, the synchronization of the EVL databases between is performed. All entries are set valid upon their creation, and their validity is

Proceedings of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)
1530-1346/03 $17.00 © 2003 IEEE

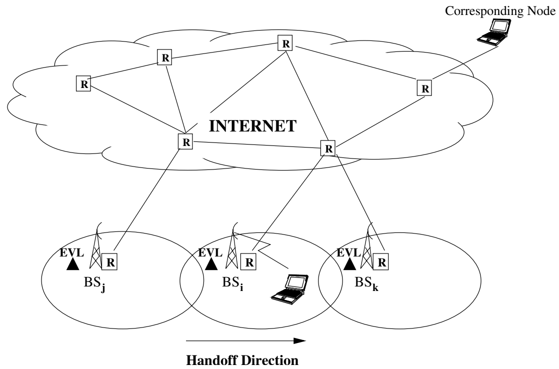

Figure 3. An illustration of wireless environment during handoff of mobile station from BS$_1$ to BS$_2$.

preserved until the EVL processors receives any signal indicating the change in network conditions which might invalidate the current acceptance status of the entry. If any resource status change event occurs in any of router resource capacity for either increase or decrease cases, i.e., invalidator event, the valid EVL entries will be invalidated to reflect the changed environment conditions to their admission control decisions. If an event is a capacity decrease of a path or a router along the path, then EVL entries, whose connection pass over that path, with accept decisions will be invalidated. In the same manner, if an event represent a capacity increase, then valid EVL entries with reject decisions will be invalidated. The invalidated EVL entries will be reprocessed and their preparation results and call admission control decision will be re-evaluated and stored again in the proper fields to be used for handoff requests.

2.3 Virtual QoS Provisioning

Decision (Accept/Reject) field evaluation of an EVL entry is a more involved task than the call admission control process of handoff requesting mobile station. Although the consideration for scarce wireless resources is still important, QoS provisioning, i.e., accept/reject state of an EVL entry, must be performed in an end-to-end fashion. Since resource requirements and expected constraints are to be provided throughout the path, the accept/reject evaluation of the EVL entries should be performed by considering the resource availability states of the routers which are on the path for connections of the incoming mobile. This will be performed as if new RSVP reservation setup is being performed by the mobile station as it is described in [6]. Although RSVP has its diagnostic messaging functionality, [8] specifically identifies checking the resource availability along the path as useful but non-goal for the diagnostic messages. However this function can be easily adopted by an extension to RSVP for resource availability query. The 8-bit *Msg Type* field in RSVP message is used for distinguishing seven different RSVP message types. This means that we can further identify two additional message types with the unused bits in that field. These two new messages are *PathQuery* and *ResvQuery* which will be used for virtual QoS provisioning.

Having a functionality to obtain resource availability status without actual reservation is indeed very useful in highly dynamic environments like wireless networks. Resource reservation for mobile stations before handoff occurs may lead to underutilization of network resources. Keeping EVL entries with the QoS provisioning results will help to reduce handoff latency without reserving network resources in advance. In order to perform this provisioning, EVL processor sends PathQuery message along the path to the corresponding node. This continues down to receiver and it sends back ResvQuery message. Here, unlike the actual RSVP setup process no resource reservation takes place. If any router can not provide the expected requirements of the request, then it will return a notification of failure. Then the decision field for this EVL entry will be set as reject. If no error is received and the ResvQuery message is received, then the decision can be set accept.

This method also reduces overhead due to RSVP signaling between two nodes. EVL processor at the base stations will do the signaling on behalf of the mobile station. By this way, signaling is performed only in wired portion of the end-to-end connection path, and the scarce wireless resources are saved.

The method also checks for resource availability of the cell in terms of wireless channels. In case of excessive bandwidth requirements, if no adequate wireless resource exist then handoff request for the incoming mobile station will be inevitably denied. This QoS provisioning result is merged by the EVL processor with the wireless channel availability status, then the final decision is made.If the entry of a mobile contains Accept in its decision field, it will enjoy seamless service. Otherwise, its handoff request will not be approved. By this way, handoff call admission control is integrated with the QoS provisioning of the mobile.

This method provides virtual QoS provisioning procedure with exact resource availability information. Any resource availability change on the way to corresponding node, will be reflected to the current acceptance status of the entries. The preemption priority may also be taken into account during this virtual resource reservation and provisioning process.

Proceedings of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)
1530-1346/03 $17.00 © 2003 IEEE

3 Performance Evaluation

The efficiency of the proposed method is directly related with the rate of entries' being valid throughout the time. An EVL entry can take the advantage of proposed method if and only if its entry has valid state at the time of handoff request. Therefore, without implementing the exact mobility patterns for the mobiles and QoS expectations, traffic characteristics in a simple shadow cluster topology, handoffs can be simulated as events arriving on a predefined rate. In the same manner, invalidation events can be modeled as events arising at certain times without assigning resource usage characteristics to mobiles on an individual basis.

The essential assumptions about the network are as follows; all events arrive based on a Poisson traffic, with exponentially distributed time intervals. Inter-cell handoff rate is identical for each current cell-target cell pair. Average number of mobiles, each cell has, remains the same.

In this simulation, one central cell neighbored by adjacent cells formed a simple shadow cluster and handoff requests of mobiles in the cluster are processed according to the handoff and invalidation event rates. The set of all invalidator events that invalidate EVL entries with accept and reject decisions are grouped into Accept Invalidators(AI) and Reject Invalidators(RI), respectively. In each cell, events arise with rates λ_{HO}, λ_{AI} and λ_{RI} and therefore entries are invalidated if necessary. EVL processors work to keep the contents of the entries ready all the time.

The average probability of an invalid entry leaves the processor with accept decision is assumed to be deterministic with ρ_A. The event rates for accept invalidator and reject invalidator arrivals are basic network parameters affecting the performance of the proposed method. On the other hand, the entry processing time(T_E), which is assumed be deterministic for simulation purposes in absence of disturbing invalidator event during a service of a group of entry, computational power of the processor(c) which determines the size of a group of entry that are processed in parallel, are two other important parameters for performance of our method. Exploring the effect of these parameters on the percentage of valid entries, which reflects the probability of the handoff requests that could take the advantage of the methodology, will give insight about the performance of the proposal under different circumstances.

Performance of the methodology has been observed at various rates of λ_{AI} and λ_{RI}, EVL processing time T_E, i.e. the average time needed to validate an entry or EVL processor computational power c, at average number of EVL entries per cell is $N = 100$, average entry accept decision ratio of $\rho_A = 0.9$ and handoff rate of $\lambda_{HO} = 0.1$. Each data point is obtained through observation of 1000 successive handoff requests.

An increase in invalidation rates results in decrease in mean percentage of finding valid entries at handoff request time. Fig. 4 shows that at the same value of invalidator event arrival rates, as T_E decreases, the probability of finding an entry valid at any time considerably increases. Hence, how fast the invalidated entries are processed, determines the efficiency of the proposed methodology.

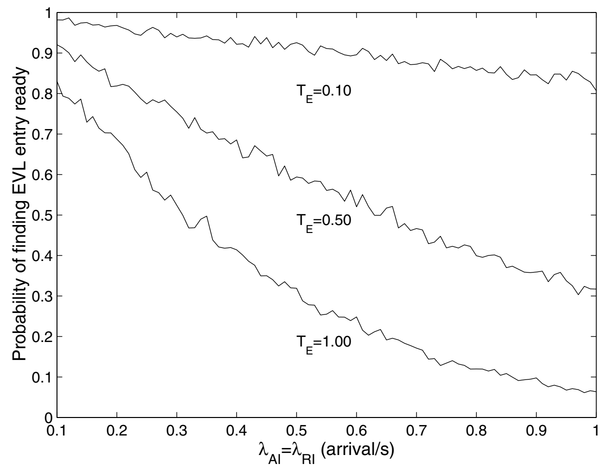

Figure 4. Probability of finding an entry valid vs. $\lambda_{AI} = \lambda_{RI}$, changing T_E= 0.1, 0.5, 1.0 for $\lambda_{HO} = 0.1$ arrivals/sec per cell-pair, $C = 50$, $N = 100$

The effect of the processor computational power, c, is seen in Fig. 5. For the maximum computational power environment, entries invalidated on arrival of any invalidation event, being the bulk input for the processor, can be directly processed and validated in at most T_E seconds. However for the limited computational power case, a queuing problem arises making invalid entries have to wait for others to be processed, giving rise to longer validation times and reduced chances of finding any entry as valid, and hence reduced chance of taking advantage of proposed method. For cases with maximum computational powers, there is no queuing problem; any invalidated entry is immediately served and validated.

Signaling overhead incurred by EVL employment scheme is shown in Fig. 6. In this case, the overhead considered includes the signaling between the EVL processors to inform the neighbor cells for the prospective mobile station and the signaling for the virtual QoS provisioning method. Overhead is represented with percentage of the assumed capacity of the wired links in network. At low and moderate invalidation rates, overhead is sufficiently low being under 10%. Overhead increases with increasing invalidation rate as EVL processors are forced to perform CAC repeatedly. However, as signaling takes place on wired links

Proceedings of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)
1530-1346/03 $17.00 © 2003 IEEE

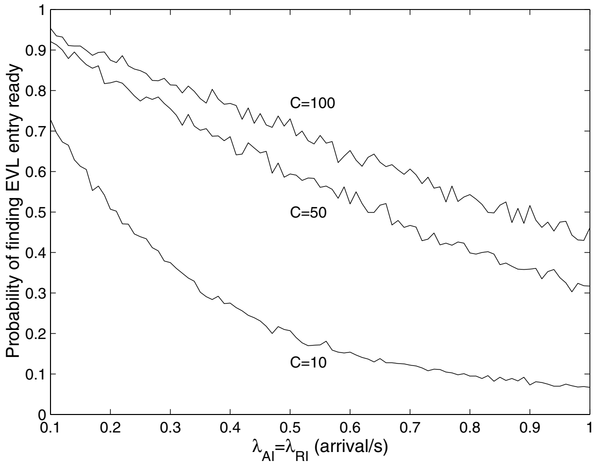

Figure 5. Probability of finding an entry valid vs. $\lambda_{AI} = \lambda_{RI}$, changing C= 10, 50, 100 for $\lambda_{HO} = 0.1$ arrivals/sec per cell-pair, $T_E = 0.5$, $N = 100$

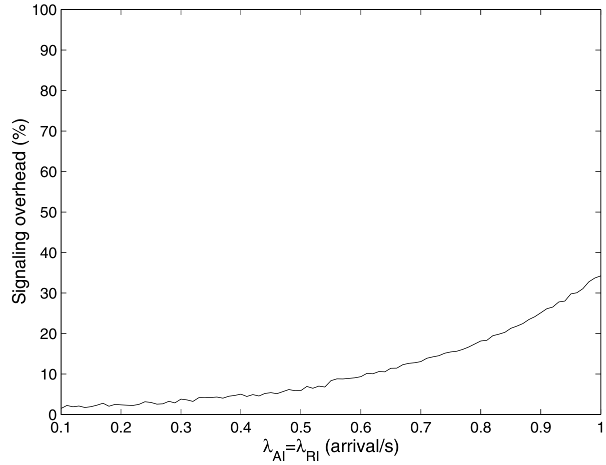

Figure 6. Signaling overhead due to EVL employment for $\lambda_{AI} = \lambda_{RI}$ for $\lambda_{HO} = 0.1$ arrivals/sec per cell-pair, $T_{EVL} = 0.5$, $C = 50$ and $N = 100$

instead of wireless channel, the cost of this overhead can be tolerable.

At moderate levels of rates of invalidation occurrences, waiting time for the handoff preparations can be improved effectively with very low signaling overhead. For the cases explored, methodology proves to be useful over 60% in a high invalidation rates environment with low overhead of $\%10$ for moderate $T_E(0.50)$ and $c(50)$ values. If invalidation events arrival rate or T_E is not sufficiently small, validation process of an entry is subject to many interruptions, by the arrival of invalidation event. This results in reduced performance of the methodology and hence increased signaling overhead. With powerful processors capable of computing entries of all the mobiles from which it is responsible, even at very high invalidation rates, handoff events of mobile stations may take the advantage of the proposal.

4 Conclusions

In this research, we have presented a new handoff method to provide QoS continuity during handoffs in next generation wireless networks. QoS expectations of the mobiles are learned in advance by pre-handoff messaging, and this information is transferred to candidate base stations. Mobile specific information is stored as an EVL entry at EVL processors. The EVL processors then perform virtual Qos provisioning with new functionality added version of RSVP in an end-to-end fashion. The results are also stored in EVL entry to be used at the actual handoff time. Any change in the resource availability will be reflected to the provision result and the entry will be kept updated to be ready in case of actual handoff. The performance simulations shows that the proposed method is a promising methodology for QoS maintenance in next generation wireless networks.

The increase in network traffic due to messaging between EVL processors seems to be a major drawback of the proposal at first glance. In order to keep EVL states updated or to perform QoS provisioning tasks, there will be considerable amount of signaling traffic. As shown in performance evaluations, at low and moderate invalidation rates, overhead is sufficiently low being under 10%. Overhead increases with increasing invalidation rate as EVL processors are forced to perform CAC repeatedly. However, as signaling takes place on wired links instead of wireless channel, due to low cost dedicated lines between access points and switching centers, this increased network traffic might be neglected compared to the cost of the saved wireless resources.

References

[1] A. K. Talukdar, B. R. Badrinath, and A. Acharya, "MRSVP:A Resource Reservation Protocol for an Integrated Services Network with Mobile Hosts," *Wireless Networks,* Vol. 7, No. 1, 2001.

[2] A. S. Acampora and M. Naghshineh, "An Architecture and Methodology for Mobile-Executed Handoff in Cel-

Proceedings of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)
1530-1346/03 $17.00 © 2003 IEEE

lular ATM Networks," *IEEE J. Select. Areas Commun.*, Vol. 12, No. 8, pp. 1365-1375, October 1994.

[3] W. T. Chen and L. C. Huang, "RSVP mobility support: a signaling protocol for integrated services Internet with mobile hosts," in *Proc. IEEE INFOCOM 2000*, Vol. 3, pp. 1283-1292, 2000.

[4] G. L. Grand and E. Horlait, "A Predictive End-to-End QoS Scheme In A Mobile Environment," *IEEE Symposium on Computers and Communications*, 3-5 July 2001, Tunusia.

[5] C. C. Tseng, G. C. Lee, and R. S. Liu, "HMRSVP: A Hierarchical Mobile RSVP Protocol," *International Conference on Distributed Computing Systems Workshop,* pp. 467-472, 2001.

[6] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification," *RFC 2205*, September 1997.

[7] D. Levine, I. F. Akyildiz, and M. Naghshineh, "A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks using the Shadow Cluster Concept," *IEEE/ACM Trans. Networking*, Vol. 5, No. 1, pp. 1-8, Feb. 1997.

[8] A. Terzis, B. Braden, S. Vincent, and L. Zhang, "RSVP Diagnostic Messages," *RFC 2745*, January 2000.

Proceedings of the Eighth IEEE International Symposium on Computers and Communication (ISCC'03)