Resource Allocation Protocol for Rich Media Wireless Networks

Prof. Nupur Prakash, Dr. C. S. Rai, and R. L. Ujjwal
University School of Information Technology,
Guru Guru Gobind Singh Indraprastha University,
Kashmere Gate, Delhi – 110403.
nupurprakash@rediffmail.com, csrai_ipu@yahoo.com, and ujjwal_rl@rediffmail.com

ABSTRACT
In the next generation high-speed wireless networks are expected to support rich media applications (video, voice, and data). As such, it is important that these networks provide quality-of-service (QoS) guarantees. Unlike wired networks, communication entities in wireless networks change their connectivity via handoff when user moves from one cell to another. Once a connection is admitted into the network, resources must be allocated at the negotiated level, for the duration of the connection. It is important to realize that, in a cellular network where the user may move through the network traversing a sequence of cells, this commitment cannot be only local to the cell in which the connection originated. If the connection is to be maintained after the user crosses the boundary between neighboring cells, the network must guarantee an appropriate level of resources in each new cell that the user traverses. In this paper a new resource allocation protocol for rich media wireless networks is proposed that uses a combination of bandwidth reservation and bandwidth borrowing to provide QoS in terms of guaranteed bandwidth, call blocking, and call dropping probabilities.

KEYWORDS
Rich media, wireless network, cell, handover, call blocking, and call dropping probabilities, bandwidth reservation.

INTRODUCTION
The wireless networks to support both data and real-time interactive multimedia traffic will be highly in demand in next generation. In the face of this more complex traffic mix, where each connection may have different requirements, providing QoS guarantees poses a difficult challenge for network providers. While best-effort service may be sufficient for the data traffic produced by simple email and web applications, newer real-time interactive multimedia applications require more stringent QoS guarantees. Admission control and bandwidth allocation schemes can help provide these guarantees in wireline networks, however the problem is much more complex due to bandwidth limitations and host mobility in the wireless networks [1, 2].

For example, in a cellular network, a host may begin a connection and negotiate a set of QoS parameters in a cell that can easily meet its requirements, but then move to a cell whose available resources are much more limited. This new cell may not be able to guarantee the same QoS that was negotiated in the originating cell. Under these circumstances, it is very hard to provide global QoS guarantees. Admission control refers to the task of deciding whether or not a certain connection request will be admitted into and supported by the network. Admission control is essential for real-time multimedia applications. These types of applications are known to specify a lower bound on network resources, below which they cannot function. If the network cannot support this minimum need, the connection request must be denied. A new connection that is denied access into the network is said to be blocked. The connection or call blocking probability (CBP) is an important QoS parameter in wireless networks. The connection or call dropping probability (CDP) is the likelihood that the network will deny resources to a connection after it is in progress. A connection may be dropped during a handoff, when the host moves from a cell with ample resources into a cell that is too congested to support more traffic. It is well known and obvious that, from the end user’s perspective, it is far less acceptable to have a call dropped than blocked. Consequently, minimizing the CDP is one of the main goals of QoS provisioning in wireless networks.

RESOURCE ALLOCATION PROTOCOL
We are witnessing an unprecedented demand for wireless network to support both data and real-time multimedia traffic. While best-effort service suffices for datagram traffic, the usability of real-time interactive multimedia applications is vastly improved if the underlying network can guarantee adequate quality of service (QoS). Admission control and bandwidth allocation schemes can offer wire line networks the ability to provide these guarantees. Due to host mobility, scarcity of bandwidth, and an assortment of channel impairments, the QoS provisioning is far more challenging in wireless networks than that of wire line networks [3, 4]. Admission control refers to the task of deciding if a connection should be admitted into and supported by the network. Admission control is necessary for real time multimedia traffic since the amount of resources requested by these connections may not match the level of resources available at the time of connection setup. Admitting a connection into the network is tantamount to a contract between the network and the connection. On the one hand, the network guarantees that a certain level of resources will be maintained for the duration of the connection. On the other hand, the connection is expected
not to request additional resources over and above these negotiated at the time of connection setup. The agreed upon amount of resources that the network guarantees to a connection is commonly referred to as QoS. These parameters include bandwidth, latency, and jitter and some parameters are specific to wireless networks.

There are two possibilities for a new connection whose requests for resources can not be met by the network, either it will be blocked or dropped. So two more parameters of QoS are call blocking probability (CBP) and call dropping probability (CDP). CBP denotes the likelihood that a new connection request will be denied admission into the network. CDP expresses the likelihood that an existing connection will be forcibly terminated during a hand off between cells due to lack of resources in the target cell. A case of this type exist when an established connection in one cell attempts to migrate into the neighboring cell, if the new cell can not support the level of resources required by the connection i.e. the hand off is denied and the connection is dropped. The CBP and CDP together offer a good indication of a network’s quality of service in the face of mobility. An additional important parameter is the degree to which the network makes an effective used of bandwidth. This parameter is called bandwidth utilization, which expresses the ratio between the amount of bandwidth use by various applications admitted into the network and either the total bandwidth requested or the total bandwidth available [5, 6, 7].

Once a connection is admitted into the network, resources must be allocated, at the negotiated level, for the duration of the connection. It is important to realize that, in a cellular network where the user may move through the network traversing a sequence of cells, this commitment cannot be only local to the cell in which the connection originated. If the connection is to be maintained after the user crosses the boundary between neighboring cells, the network must guarantee an appropriate level of resources in each new cell that the user traverses.

**SYSTEM DESCRIPTION**

The main components of the system are Cells, Base Stations, Mobile Switching Centers, and Wireline Links. The geographic area of interest is assumed to be tiled by a collection of regular hexagons referred to as Cells. The wireless communication in a cell is supported by a base station (BS). The base stations are connected to each other by the wireline links. Several base stations are connected to a mobile switching center (MSC) that acts as a gateway from the cellular network to existing wireline networks, the Internet, and the PSTN. With the active participation of the mobile hosts and the MSC, the base stations are instrumental in initiating and finalizing hand-offs [8].

Fig.1 Architecture of cellular network

The mobile hosts in a cell communicate directly with the corresponding BS that has the responsibility of handing all demands for service originated in the cell. The BS is in charge of negotiating QoS parameters, of performing admission control, and of reserving resources for on going connections. This could mean denying access to the new connections in order to provide an acceptable level of service to active connections. Every new connection would like to be accepted regardless of the current demand. The base station must be able to manage the load efficiently to permit continual cellular network traffic. From the end user’s perspective, an initial blocking of the connection is far more acceptable than to terminate an ongoing connection due to lack of resources during hand-off.

**CALL ADMISSION CONTROL AND RESOURCE RESERVATION SCHEME**

Unlike wired networks, communication entities in wireless networks change their connectivity via handoff when they move from one cell to another. The use of micro or pico-sized cells makes the role of handoff procedures very important in maintaining the service continuity and QoS guarantees to the multimedia applications. Due to the limited bandwidth resources in wireless multimedia system there should be

- efficient call admission control (CAC), and
- efficient resource reservation (RR) schemes in order to maintain desired QoS. CAC schemes enable the system to provide QoS to new incoming as well as existing calls [9, 10, 11, 12]. The RR scheme, such as the use of guard channels (GC), is adopted to reserve resources for certain higher priority calls. Obtaining a right balance between the two opposing criteria is a big challenge.

The idea is to increase the access probability for the higher priority calls, while ensuring high overall system efficiency, in the presence of multiple QoS classes such as priority, rate adaptivity as well as different mobility. The basic idea here is of the GC scheme, which gives preferential treatment to the handoff calls by reserving a fixed number of channels exclusively for them. However, such a scheme may lead to poor channel utilization because it decreases the handoff dropping rate at the cost of increasing the blocking rate for other users. To deal with this problem, the scheme used is a
dynamic resource reservation algorithm to efficiently estimate resources needed to be reserved for high priority calls, by using the distance information of mobile users in neighboring cells. Network for multimedia cannot always meet different QoS requirements of mobile users, due to resource constraints. Therefore, the system requires rules to decide who will receive the services according to predefined cost function(s), to avoid unwanted call blocking and handoff dropping while maximizing channel utilization. Usually, handoff calls are assigned higher priority over new calls. How to seamlessly transfer resources between cells during handoff is an important issue. For this, resource reservation and call admission schemes should be integrated with the handoff mechanism to provide more flexibility to all mobile users and better QoS guarantees for premium users.

The various call admission control schemes are there. They can be divided into two broad categories: Guard Channel schemes and Queuing Priority schemes [13].

GUARD CHANNEL SCHEMES

Some channels are reserved for handoff calls. There are four different schemes.

a) The cutoff priority scheme is to reserve a portion of channel for handoff calls; whenever a channel is released, it is returned to the common pool of channels.

b) The fractional guard channel scheme is to admit a new call with certain probability (which depends on the number of busy channels).

c) Divide all channels allocated to a cell into two groups: one for the common use for all calls and the other for handoff calls only.

d) Limit the number of new calls admitted to the network.

QUEUING PRIORITY SCHEME

In this scheme, calls are accepted whenever there are free channels. When all channels are busy, new calls are queued while handoff calls are blocked, new calls are blocked while handoff calls are queued, or all arriving calls are queued with certain rearrangements in the queue. Queuing Priority (QP) based methods follow the principle, when resources become available, one of the calls in the handoff queue is served [14]. If there are no available resources, call requests are being queued until resources are available again. QP scheme needs lot of buffers to deal with real-time multimedia traffic and sophisticated scheduling mechanism is needed to meet the QoS requirement for delay sensitive calls to guarantee that the queued data will not expire before they are transmitted.

The basic idea of GC-based admission control strategies is to reserve resources in each cell a priority to deal with handoff requests. In order to provide mobile users with continuous connectivity, a system reserves backup channels referred to as “guard channels” to provide preferential treatment to priority calls and handoff calls. In such a system, call requests with lower priority are rejected if the available resource is less than a certain threshold. GC strategies differ in the number of guard channels to be chosen by a base station. There are several ways in which one can ensure that handoff connections are given priority over new call attempts in congested cells. Some of these strategies involve:

- Channel Reservation [15].
- Queuing of hand-off requests [16, 17].
- Channel Rearrangement.

There are two approaches to resource reservation:

Fixed Reservation:

A certain percentage of available resources in a cell are permanently reserved for hand-off connections.

Statistical Reservation:

Where resources are reserved using a heuristic approach. These approaches range from allocating the maximum of the resource requirement of all connections in neighboring cells to reserving only a fraction of this amount. Numerous approaches for reserving bandwidth are there in the literature [18, 19, 20]. The traffic offer to a cellular network may be classified into two classes based upon that if it is real time or not. The two types are:

- Class I
  - Real time multimedia traffic, such as interactive audio and video.

- Class II
  - Non Real time data traffic, such as e-mail and web applications.

When a request for a connection is made to the network, the following parameters are provided:

- The traffic class (I or II).
- The desired amount of bandwidth, the level at which the application performs the best.
- The minimum acceptable amount of bandwidth, the level of bandwidth below which the service would become unacceptable to a user.

One of the significant feature of the call admission control and bandwidth reservation schemes is that in order to admit the connection, bandwidth must be allocated in the originating cell and at the same time reserved for the connection in all the neighboring cells. For a new connection to be admitted in a cell, the cell must be able to allocate to the connection in the desired bandwidth. For Class I connections, the call will be blocked unless the desired bandwidth can be allocated to it in the original cell and some bandwidth can be reserved for it in each of its six neighboring cells.

During a handoff an established Class I connection is dropped if its minimum bandwidth requirement cannot be met in the new cell or if appropriate reservations cannot be made on its behalf in the new set of neighboring cells. However, Class II traffic has no minimum bandwidth requirement in the case of a handoff and a call will be continued if there is any free bandwidth available in the new cell.
Numerous approaches for reserving bandwidth are there. The scheme used in statistical based on number of connections in neighboring cells, the size of connections in neighboring cells, the predicted movement of mobile hosts, and combination of these factors. In this scheme the dropping probability of Class I connections very low since the mobile hosts should find bandwidth reserved for it, regardless of the cell to which it moves. But, bandwidth may be wasted in the neighboring cells (the host can move only to one neighbor) and the blocking probability in those cells may increase because unused bandwidth is being kept in reserve. In general this scheme minimize the CDP at the expense of the CBP and give Class I traffic precedence over Class II traffic [21, 22].

RESOURCE ALLOCATION SCHEMES
There are various resource allocation schemes available in the cellular networks. The two schemes, on the basis of which this protocol is based, are:

- The Rate Based Borrowing Scheme [22].
- Max-Min Fairness [23].

RATE BASED BORROWING SCHEME
It is clear that keeping a small pool of bandwidth always reserved for hand-offs yields low CDP. However, in this scheme the size of the reserved pool is not determined by requests from neighboring cells, but is fixed at a certain percentage of the total amount of bandwidth available in the cell. In this scheme there is no overhead of communication between neighboring base stations to request and release reservations. This scheme does not allow bandwidth from reserved from pool to be allocated to incoming hand-offs unless the bandwidth is needed to meet the minimum bandwidth requirements of the connection. This scheme also gives precedence to Class I connection. Class II traffic does not make use of reserved bandwidth. In order to lower the CDP this scheme allows the borrowing resources from existing connections especially bandwidth. The important points to note about this scheme is that

1. This scheme guarantees that the bandwidth allocated to a real time connection never drops beyond the minimum bandwidth requirement specified by the connection at call setup time. This is very critical to ensuring that the corresponding application can still function at an acceptable level.
2. This scheme also guarantees that if bandwidth is borrowed from a connection, it is borrowed in small increments, allowing time for application level adaptation.
3. This borrowing scheme is also fair in the sense that if bandwidth is borrowed from one connection, it is also borrowed from the existing connections. If borrowing is necessary in order to accommodate a requesting connection (new or handoff), every existing connection will give up bandwidth in proportion to its tolerance to bandwidth loss. That is the reason this scheme is called rate based fair.
4. The borrowed bandwidth is returned to the degraded connections as soon as possible. Thus, the degradation in the QoS is transient and limited to a minimum.

This scheme features low call dropping probability, low call blocking probability, good bandwidth utilization and reasonable success with keeping both classes of connection operating nearly their desired bandwidth level.

MAX-MIN FAIRNESS
When the amount of bandwidth requested by the connections in a cell exceeds the total bandwidth available in the cell, it is unavoidable some of the connections will receive less than their desired amount of bandwidth [23]. An allocation is MAX-MIN fair if there is no way to give more bandwidth to a connection without decreasing the allocation of a connection of lesser or equal bandwidth. The key idea in max-min fairness is that if n connections need to partition b units of bandwidth, then each is guaranteed its equal share of b/n units. Connections that require at most their equal share are granted the desired bandwidth and are referred to as satisfied. Assume that out of n connections m are satisfied and the total bandwidth requested by the m connections is S units. The residual bandwidth

$$ R = m \times \frac{b}{n} - S $$

This residual bandwidth is partitioned between the remaining connections, each receiving $$ R/(n-m) $$ units.

PROPOSED SCHEME
The new proposed scheme is similar to the rate based borrowing scheme in the sense that it also keeps a pool of bandwidth reserved for handoffs. This reserved pool is used for only Class I handoffs, with the assumption that real time connections have more stringent QoS requirements than Class II connections. This is different from the Rate based borrowing scheme in the sense that some bandwidth is reserved dynamically in each of the neighboring cells. It reserves the bandwidth in all the neighboring cells whenever a new connection request comes to any of cell. That means whenever a connection request comes to any of the cell what it does it first checks that if the bandwidth can be reserved in all the neighboring cells in which the connection can handoff, if yes then only it will accept the connection otherwise block the connection.
The important features of this scheme are:

1. No Class I connection will ever have to give up bandwidth beyond the minimum level of bandwidth negotiated at the call time setup.
2. At the time of admission in any cell, some bandwidth is reserved dynamically in each of the neighboring cells. This reserved pool of bandwidth is used only for the handoff connections. Also, connections belonging to Class I only are allowed to use this reserved pool of bandwidth.
Resource Allocation Protocol for Rich Media Wireless Networks

3. If the cell does not have enough residual bandwidth to accommodate an incoming call, the existing connections will temporary have to give up a certain amount of bandwidth.

4. While reserving the bandwidth in the neighboring cells, a cap is implemented on the size of reserved pool. Not more than certain amount of total bandwidth can be reserved in a cell.

5. If bandwidth must be borrowed, it is borrowed in small increments to allow time for application level adaptation.

6. As soon as bandwidth becomes available due to a terminating call or to a mobile host leaving the cell, the borrowed bandwidth will be returned to the degrade connections.

7. This scheme is fair in the sense that if bandwidth is borrowed, all connections will have to give up an amount of bandwidth proportional to their tolerance to bandwidth loss.

CONCLUSION

In this paper a review of various resource allocation protocols and schemes are presented. A new resource allocation protocol for rich media wireless networks is proposed that uses a combination of bandwidth reservation and bandwidth borrowing to provide QoS in terms of guaranteed bandwidth, call blocking, and call dropping probabilities. This is different from the rate based borrowing scheme in the sense that some bandwidth is reserved dynamically in each of the neighboring cells. In this protocol, the call dropping probability of Class I connections very low since the mobile hosts finds bandwidth reserved for it, regardless of the cell to which it moves.

REFERENCES


