Resource reservation in a connectionless network

A. Eriksson
Ericsson Telecom
Dialoggatan 1, S-126 25 Stockholm, Sweden

phone: +46-8-719 2253, fax: +46-8-719 6677
e-mail: etxaeon@kk.etx.ericsson.se

Abstract

This paper describes a new signalling protocol that supports resource reservation for unicast traffic in a packet network. The key feature of the protocol is that resources can be reserved on a per connection basis without introducing connection states in the network. This is accomplished by the combination of connection state handling in the hosts and link state handling in the network. The handling of link states rather than connection states allows for a connectionless mode of operation in the network, which is attractive from a complexity and scalability point of view.

Keywords

Resource reservation, connectionless, Internet, Quality of Service, scalability

1 INTRODUCTION

The Internet Engineering Task Force is standardizing the Resource Reservation Protocol RSVP (Braden et al. 1994 and 1997). This protocol introduces connection states into the previously connectionless Internet. These states are used to store information in the network nodes about bandwidth, buffer parameters, identity and status on a per connection basis. However, the simplicity of the connectionless architecture is perceived as one of the key features of the Internet. The introduction of a connection-oriented protocol, such as RSVP, may lead to poor scalability properties. Possibly a complexity of the same magnitude as for the connection handling functions of a traditional telephony exchange must be added to an IP router that supports resource reservation.

One important objective for RSVP is the support of multicast applications where each user is able to make a separate resource reservation. Moreover, RSVP is designed to support bearer service classes with a tight control of transit delay and delay variation. These objectives necessitate a connection-oriented network.

The Ticket Protocol described in this paper is based on the assumption that the major part of the real-time traffic is generated by either two-party calls, or multi-party calls with only a small number of parties. For these cases unicast connections are sufficient. As long as the number of parties is small, the multi-party calls can be supported by a mesh of unicast connections. These assumptions imply that the network should be optimized for unicast connections. Multicast real-time traffic can then be supported by an overlay network of RSVP multicast routers which are interconnected by means of tunnels over the underlying unicast network.

The Ticket Protocol is also based on the assumption that absolute guarantees on the maximum network latency are not needed for most real-time applications. Interactive applications such as telephony and video conferencing do not require a firm upper bound on the delay, but rather a service that, with rare exceptions, offers a small delay.

The objectives for the Ticket Protocol are certainly more relaxed than for RSVP. As a result, the Ticket Protocol can operate over a connectionless network as described in this paper. The connectionless mode of operation is a major simplification compared to the connection-oriented RSVP. However, this simplification also implies some limitations, for example with regard to routing, packet scheduling and policing. These limitations will also be described in the paper.

2 DESCRIPTION OF THE TICKET PROTOCOL

2.1 Overview

The Ticket Protocol addresses the problem of offering traffic contracts with a QoS better than best-effort over a wide area connectionless network. Since no connection identities can be stored in a connectionless network, the service differentiation is based on the use of priority bits in the IP header. However, in a public connectionless network, there is a problem of controlling the amount of traffic using the high priority levels. Possibly everyone could be using the highest priority, resulting in no improvement compared to the best-effort service. To avoid this problem, the usage of the high priority levels must be controlled by the network. By limiting the aggregate bandwidth of the high priority traffic to a fraction of the total bandwidth on every link, a controlled QoS can be achieved.

Before a connection can use a specific priority level and bandwidth, a traffic contract is set up. This is done by means of a resource reservation request from the user that must pass admission control in the network. Traditionally the handling of the admission control and the traffic contract would be based on connection states in the network. However, it is desirable to retain the simplicity of a connectionless network. This is achieved according to the following description of the Ticket Protocol.

When initiating a unicast connection with a controlled QoS, the source sends a message to the network with a request for a specific traffic contract, i.e. a permission for a specific source to use a specific priority level with a specific bandwidth to a specific destination during a specific time. This request message is routed across the network and is subject to admission control at every router and its associated output link, see Figure 1. If the admission control is successful, the request will reach the destination host; otherwise it will be dropped. The destination host returns the request to the access router at the source. The access router recognizes that the request for a traffic contract has passed the admission control successfully, and translates it into a so called ticket message, which is sent to the source. This message contains all data about the traffic contract.

Since there may be an incentive for the user to forge the ticket message in order to get access to more bandwidth or a higher priority level than admitted in the traffic contract, the information in the ticket message is protected by the network with a digital signature. The mechanism described in (Atkinson, 1995) and (Braden et al. June 1994) can be used for this purpose.

The sender periodically transmits the ticket message to the network by inserting it in the user data packet flow. The ticket message follows the same end-to-end path across the network as the user data. The network can thus use the ticket message to extract all the information that is needed about the traffic contract of the connection (e.g. bandwidth, priority level, QoS parameters, time of expiry). Therefore, the network does not have to store a copy of the traffic contract and can operate in a connectionless mode.

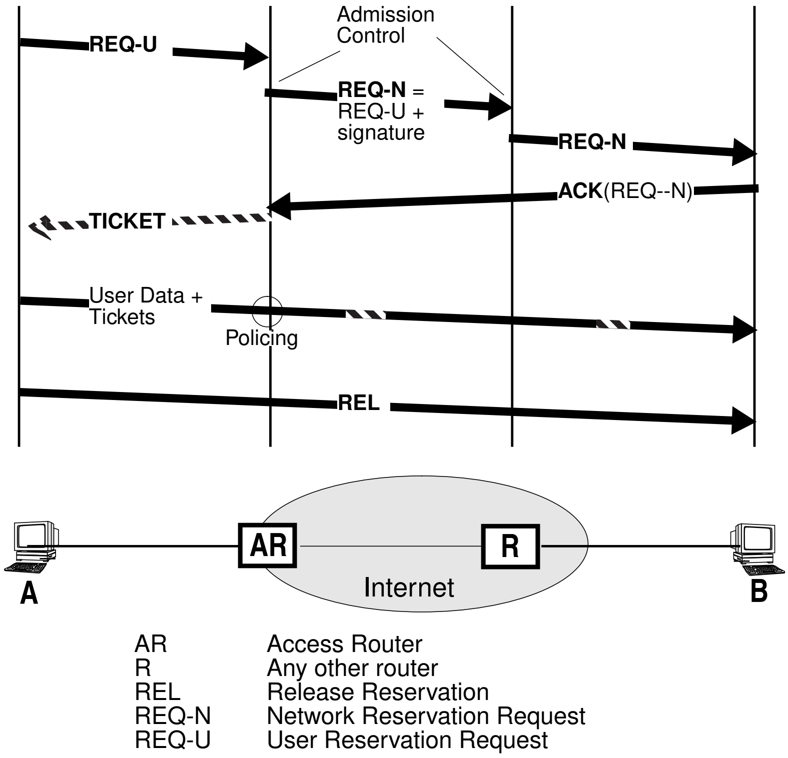


Figure 1 Overview of the Ticket Protocol signalling messages.

The network uses the information in the periodically recurring ticket messages to calculate the aggregate amount of resources that have been reserved per priority level and per link. This information is used by the admission control when deciding if a new resource reservation request should be accepted or rejected. The calculation of the aggregate amount of reserved resources requires link states, but not connection states.

The information in the ticket message is also used when specific connections are policed. For example, connections that are using a specific priority level and bandwidth without including a ticket message with a permission to use these resources should be dropped at the edge of the network.

Policing requires that a network state machine is set up for the policed connection. If policing is done on a sample basis, the number of state machines will be small. However, if policing of all connections is desired, the edge router must have a state machine per connection. The edge router then uses the Ticket Protocol in a connection-oriented mode, see section 2.5, while the core network uses the Ticket Protocol in a connectionless mode.

2.2 Detailed Description of the Ticket Protocol

Functionality

The Ticket Protocol is used for signalling between the user and the access node as well as for signalling between network nodes. This means that it supports the same type of functions as RSVP or ATM UNI/NNI signalling, that is:

- set up of a traffic contract between user and network for a specific connection;
- request for and reservation of end-to-end network resources for a specific connection;
- admission control;
- providing information from the user to the network for routing, policing and charging;
- release of the resource reservation.

Operation

The operation of the Ticket Protocol is described below. The numbers in the text below are references to specific signals or events in Figure 2, which is an elaboration of Figure 1.

1 The user sends a REQ-U message to the network with a request for reservation of network resources for a connection with a specific bandwidth, priority level and destination.

2 The access node translates the REQ-U message to a REQ-N message by adding a time of expiry parameter and a digital signature. The time of expiry is needed because the reservation is always made for a limited time interval T_t with a length in the order of seconds. The digital signature is used to protect the REQ-N message from being changed by the receiving user when it is looped back to the sender.

3 The REQ-N message is routed across the network based on the destination address and priority level in the IP header. Every router along the end-to-end path performs an admission control on the outgoing link based on the information in the REQ-N message. If the requested bandwidth and priority level can be supported by the link, then resources for the connection are reserved on the link for a time period of T_t, and the REQ-N message is forwarded along the link to the next router. If resources are not available on a specific link, the admission request is rejected and the REQ-N message is discarded.

4 If the admission control is successful on all links along the path, the REQ-N message will arrive at the destination host, which will loop it back to the sender unchanged, except for the addition of an acknowledge information element. In case the receiver is not interested in a connection, there is also an option not to return the acknowledgement.

5 When the looped back message, ACK(REQ-N), reaches the access router serving the sender, the digital signature and the time of expiry are checked. If they are correct, the ACK(REQ-N) is translated to a ticket message, which is transferred to the sending host. The ticket message is protected by the network by a digital signature, so that the sending user shall not be able to code a larger bandwidth or a higher priority level than admitted by the network.

6 The ticket message is inserted in the user packet flow, and is routed along the same path as the user data across the network. It is either inserted in every user data packet, or sent as a separate packet with a period of T_t.

7 The access router checks the digital signature and the time of expiry. The access router may also use the information in the ticket message for policing of the corresponding user connection.

8 The receiver acknowledges the ticket message in the same way as the REQ-N message.

9 When the access router receives the ACK(Ticket) message, a new ticket message is issued every period T_t by the access router with a new time of expiry, which is the value of the current time of expiry parameter plus T_t. The digital signature is recalculated taking the new time of expiry into account.

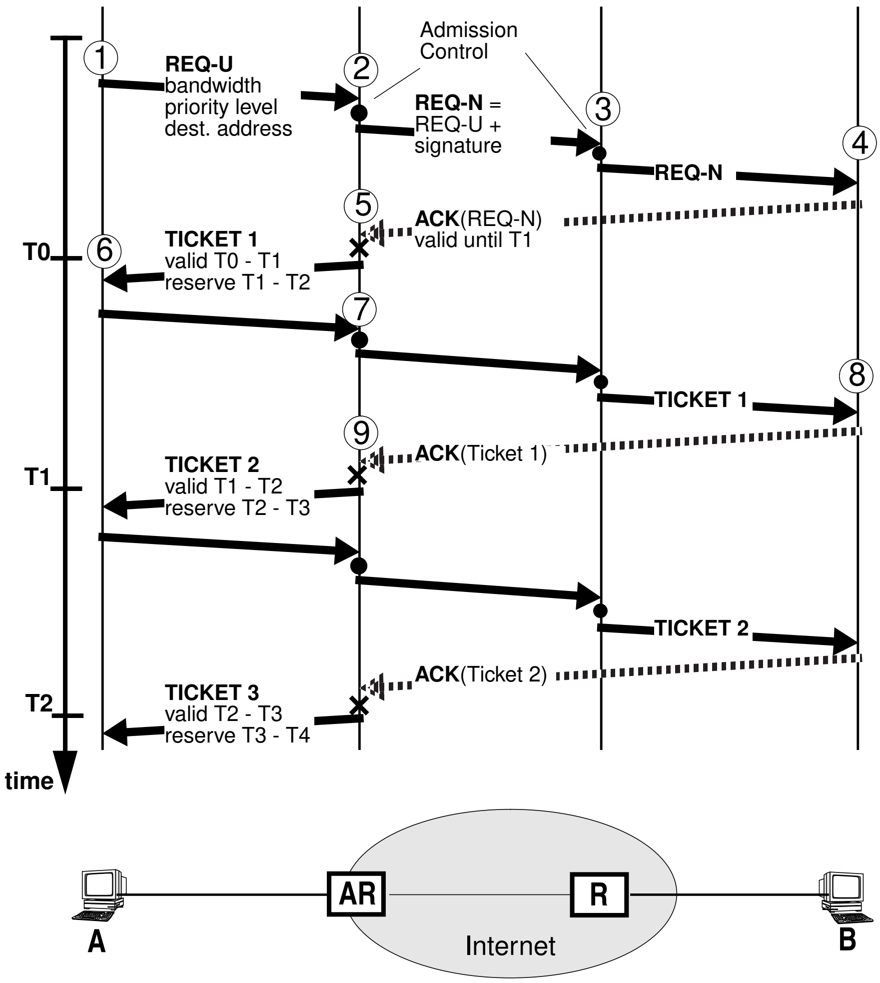
By the cyclic renewal of the ticket based on the acknowledgement of the old ticket from the receiver, a ticket loop is formed. By means of this ticket loop, network resources are reserved, even if user data are temporarily not sent. The ticket loop thus supports a per connection reservation of network resources, even though all per connection states are kept in the hosts. By inspection of the signal flow it can be confirmed that the network states are related to the aggregate reserved bandwidth per priority level and per link, and that there are no per connection states in the network.

Please note the two-fold function of the ticket message in Figure 2. A ticket message valid for a time period T1 - T2 is used both to prove that access has been admitted for that period, and also to reserve resources for the next time period T2 - T3.

The ticket message is used by the routers along the path as a source of information about the parameters of a connection, such as bandwidth, token bucket rate, source, destination and priority level. The network nodes thus do not need to maintain states for every connection, thereby being able to operate in a connectionless mode. However, also a connection-oriented mode of operation can be supported, see section 2.5.

When the sender or receiver wishes to terminate the reservation, they can do so by discarding the ticket message. The ticket loop is then broken, and no new tickets are issued. Due to the termination of the ticket loop for a connection, the links along the path of the connection will calculate a decrease in the reserved bandwidth, and thus they will have more bandwidth available for new resource reservations.

Old tickets cannot be reused due to the time of expiry parameter. Resources for a new connection can only be reserved by issuing a new REQ-U message.



AR Access Router
R Any other router
REQ-N Network Reservation Request
REQ-U User Reservation Request
✕ issue of the next ticket
● reservation for the next time slot

Figure 2 Signal diagram for the Ticket Protocol.

The ticket loop
When performing admission control, the aggregate reserved bandwidth for all connections with a specific priority level on a link is calculated during each consecutive time interval T_t. This is done by addition of the relevant parameters in the ticket messages of the active connections. The ticket message must therefore be sent with a period T_t. Also, in order to avoid that a ticket message is used repeatedly, it must be valid only for one period. The time of expiry parameter in the ticket message must thus be renewed with a period of T_t. The renewal is done by the access router.

If a user fails to send a ticket message during a time interval, then the ticket loop is broken and the reservation is released. The reason is that a missing ticket message for a connection means that the admission control function cannot take the bandwidth of that connection into account, and may grant this bandwidth to an other connection making a reservation request. The sender can check that the ticket loop is not broken by monitoring that a new ticket message for the subsequent time interval is received from the access node. If no new ticket message is received by the sender within a time interval $T << T_t$ after sending the previous one, then the previous ticket message must be resent to request a new ticket.

However, what happens if a ticket message is lost half-way along the path? When retransmitting a ticket message, some nodes will count this message twice. This shows that the estimate of the aggregate reserved resources cannot be based only on the ticket messages. Also the REQ message and the message for the release of reservations must be taken into account. Moreover, the estimate can be improved by measurement of the high priority traffic on the link, see chapter 2.3.

Release of reservations
Reservations can be released by stopping sending tickets. The ticket loop will then be broken, and the estimate of reserved bandwidth made along the path of the connection will be decreased by an amount corresponding to the bandwidth of the released reservation. This will be true either if the bandwidth estimate is based on addition of the bandwidth in the ticket messages, or if it is based on the measurement of the aggregate high-priority traffic. In the first case the decrease of the bandwidth estimate will be done within a time period T_t, while in the second case it will take a longer time, since the measurement based estimation requires averaging.

The time of the reservations release can be decreased by means of an explicit release message issued by the sender when the ticket loop is broken.

2.3 Admission control

Admission control is performed link by link based on the information in the REQ message. The admission control procedure takes the bandwidth, token bucket parameters and priority level in the REQ message into account and makes an assessment whether resources can be reserved for the requested new connection while still fulfilling the service contracts for the already admitted connections. This is done by each router on the outgoing link, and a ticket is only issued if the REQ message passes the admission control on all links along the path.

In order to determine if a new connection with a specific priority level can be admitted on a link, the aggregate resources already reserved for the connections using that priority level must be estimated. In a connection-oriented network, the aggregate resources would simply be calculated by summation of the connection parameters stored in the network. In a connectionless network there are by definition no such parameters stored in the network. Therefore the following methods can be used:

- The aggregate reserved resources for a priority level are estimated by summation of the bandwidth and token bucket parameters obtained from the ticket messages for each connection on the link.
- The aggregate reserved resources for each priority level are estimated by measurement of the traffic on the link (Jamin et al. 1997).

The first method gives a more accurate estimate of the reserved resources, since explicit connection parameters, such as peak bit rate, are available in the ticket messages. However, these messages may get lost in the network, and the second method could therefore be useful as a complement.

2.4 Policing

The policing function checks that a connection adheres to the traffic contract in the ticket message. Also, the integrity of this message is checked by means of the digital signature.

In order to police all connections continuously, state information such as token bucket parameters must be installed for every connection in the access node. This means that the access node would operate in a connection-oriented manner. If a fully connectionless network is preferred, then the policing must be done on a sample basis, i.e. a fraction of the connections are picked out for policing. The criterion for picking out a connection for policing could be pretty much the same as in an ordinary customs check, i.e. on a random basis or when an anomaly is detected.

2.5 Connection-oriented operation

The Ticket Protocol can be used in a network where some subnets are connectionless, and some are connection-oriented. The ticket message contains the complete traffic contract, including bandwidth parameters, priority level, time of expiry and destination address. This information is sufficient to set up connection states.

2.6 Handling of the Controlled-Load service

To achieve QoS differentiation, the Ticket Protocol relies on the use of priority bits in the IP header. The content of the header is the only available information when scheduling packets in a connectionless network. Scheduling mechanisms that rely on additional information, such as weighted fair queueing, can not be used on a per connection basis in a connectionless network. As a consequence, priority scheduling must be used in a Ticket Protocol network.

The network latency and packet loss rate provided by a simple priority scheduling mechanism depends strongly on the load of the high priority traffic. The definition of the Controlled-Load service specified by the IETF (Wroclawski, 1997) can be fulfilled if this load is kept below a certain level. The admission control mechanism and the policing mechanism must therefore limit the load of the high priority traffic below this level.

3 QUALITY OF SERVICE AND TYPE OF SERVICE ROUTING

The Ticket Protocol is able to support connectionless operation as well as connection-oriented. These two modes of operation are handled very differently from a routing point of view.

3.1 QoS routing in a connection-oriented network

In a connection-oriented network, each connection can be routed separately based on parameters signalled at connection setup, such as bandwidth and delay requirements. The load conditions of the network are also taken into account when making the routing decision. The QoS for the connection as well as the network utilization can thus be optimized. For example, if a link along the primary path selected by the routing protocol is congested, the routing protocol can select an alternate path. This reduces the blocking probability and improves the network utilization.

3.2 Type of Service routing in a connectionless network

In a connectionless best-effort network, the routing is normally only based on the destination address. In order to support an improved QoS, additional information such as the Type of Service (ToS) bits in the IPv4 packet header can also be used. These four bits are used to inform the routing protocol that the routing decision should optimize either for low delay, high bandwidth, high reliability, or low monetary cost (Almquist, 1992).

For parallel packet flows with identical source and destination addresses, only the ToS and precedence bits can be used to differentiate the routing in a connectionless network. As a consequence, the routing protocol cannot select an alternate path if a link along the primary path cannot support the requested bandwidth and QoS. In this case a reservation request must be rejected by the admission control mechanism. This will limit the performance of the Ticket Protocol when used in combination with ToS routing in a connectionless network.

3.3 Handling of route changes

The routing tables are updated quite frequently, for example due to routine traffic management procedures, a change in the network topology, or a link failure.

In a connectionless network, a router immediately reroutes all the traffic related to a specific entry in the routing table when that entry is updated. This works for best-effort traffic but is not allowed for already established connections with reserved resources, which first must pass an admission control along the new path. Therefore a mechanism must be introduced to prevent this immediate rerouting of traffic with reserved resources. The following mechanism is proposed.

Prior to the replacement of an output link in a ToS routing table, the ToS traffic on the link is stopped by discarding all tickets, thus breaking the ticket loop. Moreover, the priority bits in the rerouted packets are reset to a best-effort value, so that high priority connections are not rerouted along a new path, on which admission control has not been passed. The router which has made the rerouting continues to reset the priority bits until the ticket loop has been broken and the reservation thus has been released. The user is thereafter only allowed to send best-effort packets and must initiate a new reservation to obtain permission to send high priority packets.

The need for the user to initiate a new reservation after a path change is of course a limitation. However, this limitation exists also in most connection-oriented networks, e.g. the PSTN. If a failure occurs in the PSTN that requires updates of the routing tables, already established connections along a failing route must be re-established along a new route.

Re-establishment of the connection by the user may be sufficient if route changes are rare. However, if route changes are made several times per day, which is the case in some networks (Paxson, 1997), then a mechanism is needed to handle the route change without intervention by the user.

4 RELATED WORK

The Scalable Resource Reservation Protocol SRP (Almesberger, 1997) is designed to be independent of connection states in the network, as the Ticket Protocol. However, a major difference is that the Ticket Protocol passes the connection parameters to the network in explicit messages, while the connection parameters (e.g. bandwidth) are implied in the characteristics of the user data flow in the SRP. The availability of explicit connection parameters in the Ticket Protocol facilitates admission control and the operation in a connection-oriented mode as an alternative to the connectionless mode. Moreover, the explicit connection parameters in combination with the digital signature facilitates policing.

5 CONCLUSION

A new resource reservation protocol, the so called Ticket Protocol, for the support of a controlled QoS over a connectionless network has been described. As shown in the paper, this can be achieved in a simplistic and scalable manner without the complexity of a connection-oriented network. Key features and limitations have been discussed.

6 REFERENCES

Almesberger, W.; Le Boudec, J.; Ferrari T. (1997) Scalable Resource Reservation for the Internet, *IEEE Protocols for Multimedia Systems: Multimedia Networking '97.*

Almquist, P. (1992) Type of Service in the Internet Protocol Suite, *IETF RFC 1349.*

Atkinson, R. (1995) Security Architecture for the Internet Protocol, IETF *RFC 1825.*

Braden, R.; Clark, D.; Schenker, S. (1994) Integrated Services in the Internet Architecture: An Overview, *IETF RFC 1633.*

Braden, R.; Clark, D.; Crocker, S.; Huitema, C. (June 1994) Report of IAB Workshop on Security in the Internet Architecture, *IETF RFC 1636.*

Braden, R.; Zhang, L.; Berson, S.; Herzog, S.; Jamin, S. (1997) Resource Reservation Protocol (RSVP) - Functional Specification, *IETF RFC 2205.*

Jamin, S.; Schenker, S.; Danzig, P. (1997) Comparison of Measurement-based Admission Control Algorithms for Controlled-Load Service, *INFOCOM'97.*

Paxson, V. (1997) End-to-End Routing Behaviour in the Internet, *SIGCOMM'97.*

Wroclawski, J. (1997) Specification of the Controlled-Load Network Element Service, *IETF RFC 2211.*

7 BIOGRAPHY

Anders Eriksson received his Master of Science degree in 1979 from the Royal Institute of Technology in Stockholm, Sweden. In the same year he joined Ellemtel AB to work on N-ISDN prototype development. In 1987 he joined Ericsson Telecom where he has been active in various areas, including ATM switching and IP routing. He is currently working on IP traffic management.