The Interplay between Call Flow Dynamics and the Dissemination of QoS Routing Updates

Rose P. Tsang
Sandia National Laboratories
Livermore, CA 94550

1 Introduction

Network support for traffic flows with Quality of Service (QoS) guarantees usually requires a connection establishment procedure which will select a route, perform admission control and reserve resources along the selected route. If sufficient resources are not available to satisfy the flow’s QoS requirements, the flow will be blocked. In order to decrease the probability that a flow will be blocked (commonly referred to as call blocking probability), the selection of a route with the appropriate amount of available resources is very important. Since the resource availability at each network switch continually changes with the establishing and releasing of flows, network switches which perform path selection must be aware of the current resource availability, i.e., QoS state, at each switch in the network. In link state routing algorithms, information on the state of individual links is exchanged among all switches so that each switch has a complete view of the network. In order to support QoS-based routing, QoS state information must also be passed along with the link state information. Examples of protocols which provide QoS enhancements to link state routing include the ATM Forum’s Private Network-Network Interface (PNNI) [1] protocol for ATM networks and QoS Extensions to Open Shortest Path First protocols [2, 3] for IP networks.

An important issue in QoS routing is that QoS states at network switches are highly sensitive to flow dynamics, e.g., flow arrivals and departures. For instance, as flow activity increases, the QoS state at each switch fluctuates more rapidly. In contrast, standard routing protocols such as OSPF are not only impervious to flow dynamics, but link state information, e.g. connectivity or distance, fluctuates on a much slower time scale than QoS state updates. Thus in a QoS-based link-state routing protocol, in order to improve route selection based on a flow’s QoS requirement, it is important to maintain consistent and up-to-date QoS state information at all switches. Consequently, QoS state update messages must be exchanged more frequently, thus potentially incurring more communication and processing overhead.

In this paper, we study the interplay between flow dynamics and QoS routing through examining its impact on call blocking probability in the context of the ATM PNNI protocol. The PNNI specification consists of a routing protocol, based upon OSPF, and a signalling protocol, based upon the ITU-T’s B-ISDN signalling, i.e., Q.2931. In PNNI routing, the routing information exchanged includes link state information as well as ATM QoS state information such as maximum cell transfer delay (maxCTD), cell delay variation (CDV), and available cell rate (ACR). The exchange of routing information is done by controlled flooding. In PNNI, when a flow arrives at the entry of the network, the source switch uses its local view of the network to select a path which meets the flow’s QoS requirements. If it cannot find a suitable path, the Generic Call Admission Control (GCAC) of the source switch rejects the flow. If a suitable path is found, the flow set-up procedure is invoked and every switch along the path performs Actual Connection Admission Control (ACAC) to determine whether it has the requested resources. If not,
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
the flow is rejected. Otherwise, the resources are reserved. For very large networks, PNNI also supports recursive hierarchical routing. However, due to the additional complexity of aggregating topology as well as QoS metrics, we consider only non-hierarchical networks in this study.

Based on a simplified version of PNNI, we examine the relationship between the frequency of QoS state updates, the QoS-routing related control traffic overhead and the call blocking probability. For instance, if the source switch uses out-of-date information to select a path, a **false blocking** situation (i.e., the local routing table’s view of the network does not reflect the current increased resource availability), or a **false probing** situation (i.e., the local routing table’s view of the network does not reflect the current decreased resource availability) may occur. In addition, out-of-date information could also lead the GCAC to select **false routes**. We investigate these issues under varying load conditions and for different QoS state update intervals. In the remainder of this extended abstract, we present the specific problem formulation, describe our approach and report our simulation results and findings.

## 2 Problem Formulation

As mentioned in the previous section, this study is based upon a simplified PNNI model. In this model, we assume a network with flat and static topology where only the QoS states of network switches change. QoS state updates are periodically generated and disseminated through the network using controlled flooding. All flows are assumed to be unicast and constant-bit-rate. Each flow has two QoS requirements: cell rate and end-to-end cell transfer delay. For simplicity, we assume that the CTD at each link is constant. Thus the QoS state of each switch consists only of the fluctuating bandwidth availability, i.e., ACR.

Upon receipt of a connection request, a source switch must select a path that satisfies the request’s QoS requirements. Path selection is computed on-demand for each flow based upon the source switch’s local view of the network state. In the first step, all the links whose ACR is smaller than the requested cell rate are pruned. Then a standard Dijkstra’s shortest path algorithm is run using CTD as the cost metric to find the path with the smallest end-to-end CTD. If this delay is greater than the requested end-to-end delay, the call is rejected. Otherwise the call setup procedure is initiated, i.e., the GCAC procedure returns a success, and the flow’s request is passed through the switches along the selected path. Each switch along the path performs the ACAC procedure where either the requested resources are reserved or the requested resources cannot be reserved and a reject message is sent back to the source switch. This latter case is referred to as **crankback**. In the case of crankback, we do not attempt any alternate routing.

When call blocking occurs at the GCAC procedure, there are two possibilities: either the local view of the source switch reflects the true state of the network, thus the flow cannot be admitted (this is referred to as **true blocking**); or the local view does not reflect the current increased resource availability so the call is blocked even though there does exist a route in the network which would have met the flow’s QoS request (this is referred to as **false blocking**). When crankback occurs at an ACAC, it may be caused by either of the following two cases: (i) no path in the network exists which can support the flow’s QoS request (i.e., the GCAC should have rejected the flow if up-to-date QoS state were available), referred to as **false probing**, or (ii) there does exist a path in the network which would support the flow’s QoS but due to the out-of-date local view of the source switch, it was not selected. The latter case is known as **false routing**. Besides call blocking probability, the other performance metric we consider is the overhead generated by the QoS state updates. We will examine the tradeoffs between minimizing the discrepancy of a switch’s local view of the network with the actual state of the network and regulating the amount of control traffic in the network.

The results of the interactions between the flow dynamics, frequency of QoS state updates and the call blocking probability are not always clear. For instance, given certain flow dynamics, say in terms of flow inter-arrival time,
how does the frequency of QoS updates affect the different types of blocking? There is an obvious tradeoff between call blocking probabilities, in general, and the overhead created by the QoS state updates. But the current network load must also be considered. If the network load is very low, and the requests for bandwidth are also low, a low call blocking probability will be attained regardless of the frequency of QoS state updates. On the other hand, if the QoS state of the network fluctuates more rapidly, QoS state information must be exchanged very frequently. Otherwise the source switch will likely either under-estimate the current amount of available resources in the network, resulting in false probing, or over-estimate the current available resources, resulting in false blocking. We will address these issues in the next section.

3 Preliminary Results and Discussion

In this section, we describe our simulation setting and discuss the results. The simulation is conducted using ns2.0 [4]. We use the topology shown in Figure 1. The topology, albeit simple, has enough connectivity to allow for a reasonable number of alternate routes. Each link has 10 Mbps bandwidth. The delay on the diagonal links is 15 μs while it is 10 μs on all other links. These delay values are chosen so that some routes are preferable over others even with the same number of hops.

Flows are generated according to a Poisson process. The source and destination of a flow are chosen randomly among the eight nodes on the boundary. As mentioned in Section 2, all flows are constant-bit-rate, each requiring 1 Mbps bandwidth and an end-to-end delay of less than 45 μs. The holding time of flows is exponentially distributed with mean of 5.0 sec.

In order to investigate the interplay between flow dynamics and update intervals of QoS states, we measure the number of flows blocked due to out-of-date QoS states under varying network loads. The network load is controlled by changing the arrival rate while keeping the holding time fixed. In the simulation results we presented below, three update intervals are used: 0.5 sec, 1.0 sec and 2.0 sec. As a reference for comparison, the number of flows blocked, assuming that the switches have up-to-date states, is also measured (this is shown in Figure 2 with update interval equal to 0.0 sec).

We first consider the scenario where the flow arrival rate to the network is constant over time. In Figure 2(a), the
total call blocking probability is plotted as the flow arrival rate increases for the three update intervals (0.5 sec, 1.0 sec, 2.0 sec) as well as for the instant update intervals (0.0 sec). As expected, the call blocking probability increases as the network load increases for all the cases. In the instant update case, flows are blocked solely because routes with sufficient resources are not available. On the other hand, in the other three cases, flows may be blocked either because there are no available routes, referred to as the no route case, or because one of the links on the route selected by the source switch does not have sufficient resources at flow set-up time (i.e., crankback occurs). The percentages of these two types of blocks are shown in Figure 2(b). We can see from the figure that the significant fraction of total blocks are actually due to crankbacks. Clearly, this is caused by the out-of-date states at the source switches.

To further examine the reason for crankbacks, we measured the number of crankbacks due to false routes. In Figure 3(b), we show the false routes as a percentage of total crankbacks. Although crankbacks due to false routes decrease as the flow arrival rate increases, they still contribute a considerable portion of the total crankbacks (more
than 78% even at a high arrival rate of 25 flows/sec). This indicates that appropriate route selection in the face of out-of-date QoS states is vital in reducing the number of crankbacks. The remainder of the crankbacks, i.e., false probes, correspond to those flows that should have been rejected by the GCAC if the QoS state was up-to-date.

Out-of-date QoS state could also lead to false blocks, i.e., those flows that should have been admitted if the QoS state was up-to-date. In Figure 3(a), we show false blocks as the percentage of blocks due to no routes. From the figure, we see that larger QoS state update intervals result in higher numbers of false blocks.

From the results presented above, we observe that out-of-date QoS states can cause the GCAC at the source switch to either under-estimate or over-estimate the resources available in the network, which will lead to the occurrence of false blocks or false probes. We expect this situation to worsen if the QoS states fluctuate more rapidly. In order to induce more rapid fluctuations in the network state, we use Markov processes to modulate the flow arrival rates. We consider both a two-state model and a three-state model which reflect the network state transition among periods of various loads. The preliminary results have demonstrated that varying the network state transition among periods of various loads can cause even higher false probes and false blocks than the constant load scenario discussed above. The results are not included here due to the space limitations.

4 Conclusion and Future Work

In this paper, we studied the interplay between the flow dynamics and update intervals and its impact on the route selection. We observed that in the presence of out-of-date QoS state, the majority of blocking is due to the selection of false routes. In addition, as the network load increases or the flow dynamics fluctuate more rapidly, the number of flows blocked due to false blocks and false probes also becomes significant.

Our study shows that choosing an optimal path such as shortest delay path in the face of out-of-date QoS states may not be the best strategy. We are currently investigating alternative route selection algorithms that can potentially achieve better performance. In order to deal with rapid QoS state changes, we are also exploring efficient update mechanisms to disseminate the QoS state information. A promising scheme is to combine event-driven update with periodic update that dynamically adapts to the flow dynamics. More extensive simulation with self-similar flow traffic is also under consideration. We plan to extend our work to more complex networks with hierarchical topology aggregation.

References