HETEROGENEITY-AWARE SHORTEST PATH ROUTING: FLOW HOLDING TIME, USER DEMAND AND NETWORK STATE

SHANCHIEH YANG, XUN SU AND GUSTAVO DE VECIANA

Department of Electrical and Computer Engineering
University of Texas, Austin, TX 78712
Email: {scyang,xsu,gustavo}@ece.utexas.edu

***Abstract*— We investigate possible performance improvements by exploring heterogeneity of traffic characteristics when designing a shortest path routing scheme. First we focus on the effect of the maintenance of the link metrics for connections with different holding times. We found that by using a differentiated routing scheme with respect to connection holding times, one can enhance network performance for a range of traffic loads. Second we propose a selective routing control scheme which determines whether to accept a shortest path routing decision based on user demands in the (source,destination,bandwidth request) tuple. Simulations were conducted to exhibit the effectiveness of such routing algorithm. Finally, realizing the difference in routing "local" versus "transit" traffic, we present a novel approach to aggregate network state by drawing on an analogy to circuit theory. The proposed "effective capacity" abstraction is not only efficient in terms of the signaling savings, but also maintains a compatible routing metric to represent aggregated and local states.**

I. INTRODUCTION

A variety of challenges have been brought to the fore in dealing with routing in emerging broadband multi-service networks. Although significant efforts have been made to develop mathematical models [11], [10], [5], [2] and heuristic algorithms [8], [12] to enable quality of service (QoS) guarantees to users, many open questions remain. Among these questions are how to deal with, and make the most of, traffic and network ***heterogeneity***.

Heterogeneity arises in many aspects, *e.g.,* network design/management, flow control, call admission control, routing, etc [10]. In this paper we focus on the impact of heterogeneity on routing performance. In particular, we are interested in the effect of heterogeneity in ***flow holding time***, ***user demand***, and ***network state availability***. We propose to improve on the traditional dynamic routing schemes, *i.e.,* shortest path routing scheme based on the residual bandwidth available on links [8], by actively exploiting such heterogeneity. The goal is to show that by *intelligently* realizing/utilizing such heterogeneity one can achieve significant performance gains with minimal additional overhead.

This paper makes the following contributions. In Section III we examine routing performance when one operates with "adaptive" and "dynamic" link metrics. Observing that connections with different holding times respond differently to the use of adaptive or dynamic metrics, we propose a differentiated shortest path routing mechanism that outperforms the traditional routing schemes in terms of the percent routed volume. We then investigate the impact on performance of routing heterogeneous user demands, namely the *(source, destination, bandwidth request)* tuple. With a view on efficiently utilizing network resources, both in terms of the source-destination distance and the bandwidth requirement, we propose and evaluate a second level of routing differentiation in Section IV. Finally in Section V we consider the fact that large scale communication networks may no longer permit flows to be routed based on *exact* network state information, either due to scalability or routing overhead limits [7]. Researchers have proposed various ways to aggregate network topology accompanied by hierarchical routing to resolve this problem, see [4], [9] for examples. With the need to support a mixture of both local and transit flows which would presumably see the exact and abstracted network states respectively, we propose a novel hierarchical shortest path routing scheme based on a consistent routing metric over aggregated/abstracted resources. We validate our proposal by showing significant overhead reduction with minimal performance decrease. We will begin by stating our network model in the next section.

II. NETWORK MODEL

We model the network by a graph $\mathcal{G}(\mathcal{N}, \mathcal{L})$ where $\mathcal{N}$ denotes a set of nodes and $\mathcal{L}$ a set of links. Each link

This work is supported by National Science Foundation Career Grant NCR-9624230, and by an Intel Technology for Education 2000 equipment grant.

$l \in \mathcal{L}$ has a capacity c_l units of bandwidth. A connection request corresponds to a tuple (s, d, b, t) where s and d correspond to the source and destination nodes, b is the required/requested bandwidth and t is the connection's holding time. When deciding a path for the connection request, we use a shortest path algorithm based on link metric $1/r_l$, where r_l is the residual bandwidth of link l[8].

For our simulations, we assume user connections arrive according to Poisson processes, and the amount of bandwidth requested is distributed according to a Bimodal distribution. We shall consider exponential, Pareto, and bimodal distributions for connection holding times while maintaining the same average load. Our primary performance measure is the *percent routed volume* of traffic, i.e., the fraction of the total offered load that is routed/accepted.

III. Heterogeneity in Connection Holding Time

Link metrics given by the reciprocal of the residual bandwidth, denoted by $f_l = 1/r_l$, are often used in shortest path routing schemes that aim to achieve high resource utilization, see *e.g.*, [8]. Given this choice of metric, a key design issue is still whether to advertise the most up-to-date value for a link's state or average it over a given time scale. A "dynamic" routing scheme uses the most up-to-date f_l value, while an "adaptive" scheme uses link states that have been averaged/filtered over time.

Note that neither dynamic nor adaptive routing approach is best for all possible network instances. Now suppose we have prior knowledge concerning the mean holding times of the incoming connections, can one improve routing performance? Our comparison of dynamic and adaptive routing schemes reveals no clear winner if one applies either scheme blindly. However, we observe that two factors interact: the "build-up" of a good average load (via adaptive routing) and the ability to track the instantaneous traffic fluctuation (via dynamic routing). Based on this observation we propose a *differentiated* routing scheme, where we assume upon arrival we can classify connections as having long or short holding times. We propose to route long (holding time) connections based on filtered metrics and short ones based on the current state information. The intuition is that for the long connections the decision we make at the moment of their arrival should be good "on average" during the time of their stay. Hence it makes sense to use the smoothed-out/filtered link metrics. For short connections, however, the priority is to seek the best possible path based on current link states, and mistakes or aggressiveness in routing such connections would quickly subside.

We show a set of simulation results to exhibit the performance achieved by the adaptive, dynamic, and differentiated shortest path routing schemes. We consider the network shown in Fig.1. This is a commercial network, whose parameters, including link capacities and traffic demands are given in [3]. Fig. 2 shows the percentage of routed work (in bits) under the three policies as traffic load increases. It can be easily observed that differentiation based on flow holding times can achieve a better performance over a range of traffic loads.

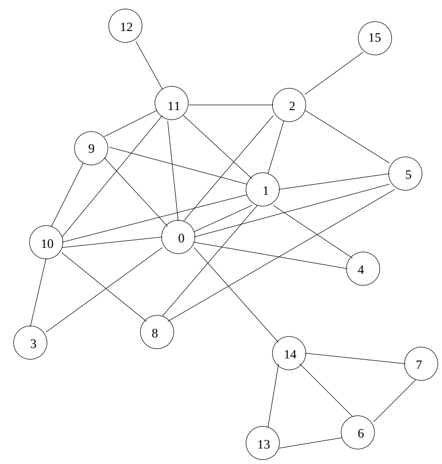

Fig. 1. The commercial network.

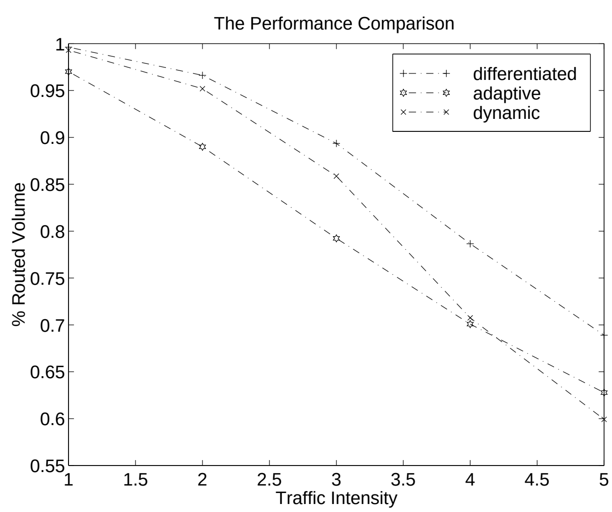

Fig. 2. Percent routed volume performance comparison of adaptive, dynamic, and differentiated shortest path routing.

IV. Heterogeneity in User Demand

Dynamic routing schemes aggressively seek paths that meet users' requests. The drawback, however, is that these paths, though apparently good for a given user, might negatively impact future arrivals. The problem usually arises from excessive occupation of network resources by individual connections which precludes more efficient use of resources in the future. A sensible routing scheme, if properly designed, tends to balance the need to accommodate connections while limiting resource consumption. Observe that we are faced with making such routing decisions for "heterogeneous" user requests. Specifically, here we focus on the differential treatment of the tuple of *(source, destination, bandwidth request)* so that "overall" routed traffic volume is enhanced.

As is well known in a fully connected single service network a single-link path provides the most efficient way of using resources. Trunk reservation techniques [1] on traditional telephone networks utilize this idea to reserve a portion of capacity on each link for future direct connections. However, this simple principle is difficult to extend to a multi-service mesh topology. We propose to incorporate two factors that impact the design of a routing scheme: (1) the size of the bandwidth request, and (2) the distance (minimum number of hops) between the demand source and destination, to preclude excessive use of network resources in the following explicit manner. Suppose that when a request enters the network, it should only be accepted if the shortest path metric for its route does not exceed that of the "worst case" it might see on the "most directed path". Assuming once again that the reciprocal of the residual bandwidth is used as the link metric, the maximum path length should a user requesting a bandwidth b be allowed to see is given by $\rho = h_{sd}/b$, where h_{sd} is the minimum hop-count from source s to destination d. The proposed routing algorithm is as follows:

1) **Selective_Routing_Control**(s, d, b):
2) **if** shortest path $p^* \in \mathcal{P}_{sd}$ satisfies:
3) $\sum_{l \in p^*} \frac{1}{r_l} \leq \frac{h_{sd}}{b}$ $\forall l \in p^*$ and $r_l > b$
4) **then** accept the request on p^*
5) **else** reject the request

We conducted simulations to verify the merits of this proposal. Fig.3 shows representative simulation results to indicate the performance gains achieved with the proposed selective routing scheme. Again we consider the commercial network shown in Fig.1. We experimented with various flow holding time distributions, *i.e.*, exponential, Pareto, and bimodal, and found similar improvement trends for all three distributions when the proposed scheme is employed. Note that the degree of performance gains achieved by our algorithm increases as the traffic intensity grows. This is consistent with arguments in previous work for telephone network where "freer" alternate routing is preferred in light loads and trunk reservation comes into play in modest to heavy load regimes.

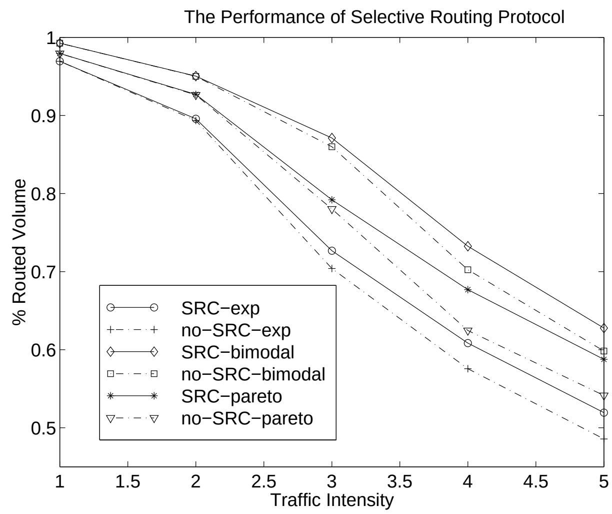

Fig. 3. Performance achieved with and without the selective routing control for heterogeneous user demands.

V. Heterogeneity in Available Network Information

With the presence of local and transit traffic flows, a mixture of routing decisions based on exact and abstracted/aggregated network states will determine how a network is loaded and thus its overall efficiency. The key to enhance overall network performance for both types of flows is to provide a consistent notion of cost metrics for abstracted/aggregated network resources. Our goal in this section is to develop such a cost metric and compatible hierarchical routing strategy for both local and transit flows so as to reduce the overhead associated with updating network state while maintaining good performance.

We consider a hierarchically arranged network where each subnetwork consists of a set of border nodes interconnecting subnetworks, and a set of interior nodes[1]. Our approach to aggregating/abstracting a subnetwork's available resource is to find a representation of "capacity availability" (shortness) between border nodes. The key abstraction of capacity availability between a border node

[1] See [4], [9], [7] for a detailed description of hierarchical network models.

pair is to find a balance among the number of possible routes connecting the pair, the number of links each route traverses, and the minimum residual capacity for each route. Note that links in series (parallel) impede (expedite) traffic transmission, and clearly larger residual capacity means higher transmission capability. Hence there is a natural analogy [2] between electrical resistance and the "transmission resistance" of a link. We view each link as having a resistance $\frac{1}{r_l}$ where r_l is the residual capacity of link l. The effective resistance between two border nodes thus can be found by mirroring the concepts in electrical networks. We define the effective capacity which represents the capacity availability of any given border node pair as the inverse of the corresponding effective resistance.

In addition to well abstracting the capacity availability, the proposed metric is consistent with the link state metric we have been using. In fact, as a hierarchical routing algorithm progresses to resolve the physical paths based on the aggregate abstraction, the transit flows naturally spread out on each subnetwork being traversed, in the same manner as would local traffic. Such consistent load balancing results in a performance comparable to that would be achieved when all flows can obtain exact network state information, while clearly hierarchical routing benefits from the saving in the state update overhead.

Two components constitute the hierarchical routing algorithm: (1) creation and maintenance of the hierarchical state aggregation, and (2) a loose source routing coupled with physical path construction. For the first part, one needs to maintain a full mesh of effective capacities between border nodes for each subnetwork. This computation is inherently distributed and local and hence scales well. For each level in the hierarchy we set up a "percentage triggering" update policy to track the dynamics of residual (effective) capacity of the (logical) links. A link state update occurs when the ratio of the change to the current state exceeds the threshold. An update at level i link triggers a test for the possible state update for logical link at level $i+1$. Clearly using a larger threshold decreases the number of update signals.

The second component consists of two conceptual steps. Given a connection request, the source node first computes a shortest path according to its aggregate image of the network. The path may contain logical links if the source and the destination nodes are not in the same subnetwork. In this case, a physical path needs to be progressively constructed by expanding each logical link.

We compare, via simulation, the performance achieved by our proposed hierarchical routing scheme to that by a flattened routing algorithm that based on actual link states. A 50% level 2 and 1% level 1 update threshold were used. The performance measures considered are again the percent routed volume and *update savings*, $\frac{U_{flat}-U_{hie}}{U_{flat}}$, where U_{flat} and U_{hie} are the average number of updates received per node, for the flattened and hierarchical algorithms respectively. We consider the two-level hierarchical network shown in Fig.4. Fig.5 exhibits a representative performance comparison in terms of both the percent routed volume and the state update savings when traffic load increases. These results reveal that our proposed scheme achieves similar performance to that by the flatten algorithm (difference $\leq$ 3%), while produce 70-80% update saving.

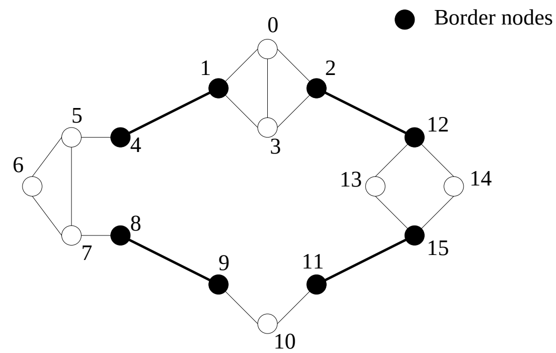

Fig. 4. A 2-level hierarchical network.

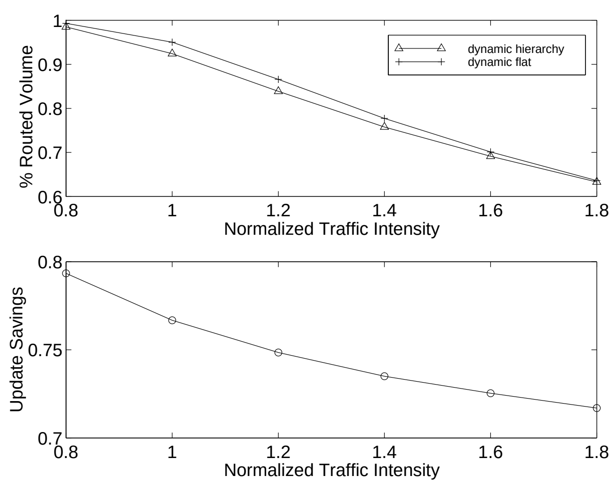

Fig. 5. Percent routed volume comparison (top) and update savings (bottom) achieved by the effective resistance hierarchical routing.

We further investigate this result by distinguishing the performance for "transit" and "local" traffic. We

[2] See [6] for other such analogies for communication networks.

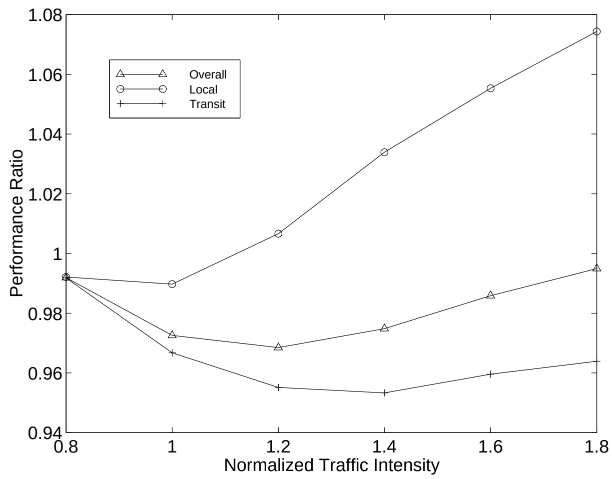

Fig. 6. Performance comparison: transit vs. local traffic.

refer to traffic which has its source and destination node in the same lower level subnetwork as local traffic and otherwise as transit. Fig.6 plots the *performance ratio* for overall, transit, and local traffic. The performance ratio is defined as the ratio of the percent accepted volume between hierarchical and flattened schemes and indicates the performance degradation (<1) or improvement (>1) when using a hierarchical scheme. As the traffic intensity increases, the performance of the transit traffic is sacrificed to boost that of the local traffic. In this example, local traffic even demonstrates a better performance when using our hierarchical scheme in heavy load regime. The reason for this is that, when hierarchical scheme is used, transit traffic does not have the same awareness of the remote network dynamics as it could have with a flattened view. Hence, local traffic has better chance to utilize its local resources. that across a wide range of traffic loads.

VI. Conclusion

In the context of shortest path routing used to satisfy user bandwidth requirements, we have explored the impacts on performance of heterogeneity in flow holding time, user demand, and network state availability. We observed that dynamic metrics are well suited to connections with short holding times, while the adaptive (filtered) ones work nicely to reflect the average network state for flows with long holding times. This motivated us to propose a differentiated routing scheme that draws merits of both metrics. To preclude excessive use of network resources, we also proposed a route selection control scheme based on heterogeneous user demands in terms of the source-destination distance and bandwidth requirement. Simulation results exhibited that the control scheme achieves better performance for a variety of traffic mixes. Our final contribution is in dealing with the coexistence of heterogeneous flows that would be routed based on both exact and abstracted/aggregated network states. Drawing an analogy on electrical networks, we proposed an "effective capacity" aggregation scheme that provides a consistent notion of "shortness" at all levels of hierarchies. Network performance achieved by employing the proposed aggregation scheme is comparable to that achieved by having exact information for all flows, while the aggregation scheme exhibits significant overhead reduction.

References

[1] R. Gibbens, F.P. Kelly, and P.B. Key. Dynamic alternative routing – modeling and behavior. *Proceedings of the 12th International Teletraffic Congress*, June 1988.
[2] A. Girard. *Routing and Dimensioning in Circuit-Switched Networks.* Addison-Wesley, 1990.
[3] A.G. Greenberg and R. Srikant. Computational techniques for accurate performance evaluation of multirate multihop communication networks. *IEEE/ACM Trans. Networking*, 5(2):266–277, 1997.
[4] F. Hao and E. W. Zegura. On scalable qos routing: Performance evaluation of topology aggregation. *IEEE INFOCOM 2000*, 2000.
[5] F.P. Kelly. Loss networks. *Ann. Appl. Prob.*, 1:317–378, 1991.
[6] F.P. Kelly. Network routing. *Proc. R. Soc. Lond. A*, 337:343–367, 1991.
[7] W. C. Lee. Topology aggregation for hierarchical routing in atm networks. *Computer Communication Review*, 25(2):82–92, 1995.
[8] Q. Ma and P. Steenkiste. On path selection for traffic with bandwidth guarantees. *IEEE INFOCOM98*, 1998.
[9] M. Montgomery and G. de Veciana. Source routing through clouds for connection oriented networks. *IEEE INFOCOM '98*, 2, 1998.
[10] K. W. Ross. *Multiservice Loss Models for Broadband Telecommunication Networks.* Springer-Verlag, Berlin, 1996.
[11] C.-F. Su and G. de Veciana. On statistical multiplexing, traffic mixes, and VP management. *IEEE INFOCOM '98*, 1998.
[12] Z. Wang and J. Crowcroft. Quality-of-service routing for supporting multimedia applications. *IEEE Journal Selec. Areas Commun.*, 14(7):1228–35, Sept. 1996.