

PERFORMANCE ANALYSIS OF PNNI ROUTING IN ATM NETWORKS: HIERARCHICAL REDUCED LOAD APPROXIMATION*

Hui Xie
John S. Baras

Electrical Engineering Department
Institute for Systems Research
University of Maryland
College Park, MD 20742

ABSTRACT

In this paper, we investigate the performance issue of PNNI routing for ATM networks. A brief introduction to PNNI routing protocol is given. We extend the reduced load approximation algorithm to ATM networks with PNNI routing. A hierarchical reduced load approximation scheme is proposed. The scheme includes two steps: the reduced load approximation algorithm is first applied on the aggregated network at upper level; then the approximation is refined at the lower level as needed using the results from the upper level. This hierarchical scheme follows the PNNI routing framework and results a fast and scaleable algorithm.

1 INTRODUCTION

The commercial communications infrastructure continues to experience dramatic growth. To modernize current military legacy communications systems, it is very important to leverage commercial leadership and expertise in communication technology arena. Recently, a program is setup to extend advanced commercial telecommunications and information distribution to a wireless, mobile battlefield environment [1]. One major trend of evolution of telecommunications networks is the rapid growth of ATM networks that are gradually covering more and more of the world. A lot of research is going on to incorporate ATM technology into wireless, mobile environment. The difference between commercial and military communication systems are the environment and performance requirements. It is crucial to have an in-depth understanding of the performance of a communication system under different environment.

In this paper, we consider the performance issue of PNNI (Private Network-to-Network Interface) routing protocol from ATM Forum. PNNI specifies a topology state routing protocol in which nodes flood quality of service (QoS) and reachability information so that all nodes obtain knowledge about the state of the network and available network resources. To reduce the complexity, PNNI uses a hierarchical model for topology

aggregation [1]. Some important performance metrics for PNNI routing are call setup delay, call blocking probability and resource utilization efficiency. However, only the call blocking probabilities for PNNI-routed ATM networks are discussed in this paper. The reduced load approximation algorithms [2][3][4] are extended to ATM network with PNNI routing. A hierarchical reduced load approximation algorithm is proposed.

Section 2 of this paper describes the reduced load approximation for general loss network. A brief introduction of PNNI routing is given in section 3. In section 4, the reduced load approximation algorithms are extended to ATM networks with PNNI routing, and the hierarchical approximation algorithm is presented. The paper concludes with section 5.

2 REDUCED LOAD APPROXIMATION FOR MULTIRATE LOSS NETWORK

Traditionally, the loss probability of a telephone system is estimated using *Erlang's formula*, which is applicable to a single link and a single call type system. Recently, a generalization of Erlang's model was proposed [2], which treats multirate integrated service networks with state dependent admission control and routing. However, the generalization leads to a combinatorial explosion in systems of equations. Indeed, the computational complexity is exponential in the number of routes between node pairs, and is exponential in the number of bandwidth classes. To circumvent the numerical complexity, researchers have developed an approximation called the *Erlang fixed point approximation* or *reduced load approximation* [3][4][5]. This technique applies to synchronous transfer mode (circuit-switched) services and to asynchronous transfer mode (ATM) services. It provides an appealing alternative and complement to discrete event simulations.

Consider a network with N nodes. A link (i, j) between nodes i and j has capacity $C_{i,j}$, counted in bandwidth units termed *trunks*. Calls offered to the network fall into S classes. A call of class s has bandwidth b_s , meaning that if the call is admitted to the network then it is allocated b_s trunks on each link of a path from its source to its destination. A link's state is

*This work was supported in part by the Advanced Telecommunications/Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAL01-96-2-002.

characterized by the number of calls in progress n_s of each call class. Consider routing policies where each node pair (i, j) is assigned an ordered list P_{ij} of paths between nodes i and j : $P_{i,j}^{(1)}, P_{i,j}^{(2)}, \dots$. The call is admitted on the first path in the list where it is admissible. If the call is not admissible on any path then it is blocked. Calls of class s are assumed to arrive to node pair (i, j) according to a Poisson process with rate $\lambda_{i,j}(s)$. If admitted to the network, the call holds for an exponential period of time with mean $\mu_{i,j}(s)$. Thus, these calls constitute a load of $\rho_{i,j}(s) = \lambda_{i,j}(s) / \mu_{i,j}(s)$. The performance metrics that we are interested in are the blocking probabilities or equivalently the admissibility probabilities. The fixed point approximation is based on assuming that a call is blocked independently by the different links along its path and that arrivals of calls of a given class to a given link are described by a Poisson process. The resulting system of equations whose unknown are, for each link (i, j) and call class s :

- arrival rates $\nu_{i,j}(s)$ of calls of class s from any node pair that includes link (i, j) on some route, given that link (i, j) is in a state that admits calls of class s , and
- probabilities $a_{i,j}(s)$ that link (i, j) is in a state that admits calls of class s .

The fixed point formulation essentially consists of two mappings. One is for every node pair (i, j) from probabilities $a_{k,l}(s)$ (with k and l dependent on the routing structure) to rates $\nu_{i,j}(s)$. The other mapping is for each link (i, j) from rates $\nu_{i,j}(s)$ to probabilities $a_{i,j}(s)$.

Let $\nu_{i,j}^{(m)}(k, l, s)$ be the rate that node pair (k, l) admits class s calls on its m th route, if that route contains link (i, j) ; otherwise $\nu_{i,j}^{(m)}(k, l, s) = 0$. Thus the aggregate arrival rate of calls of class s at link (i, j) is given by

$$\nu_{i,j}(s) = \sum_m \sum_{k,l} \nu_{i,j}^{(m)}(k, l, s) \quad (1)$$

and by the link independence assumption,

$$\nu_{i,j}^{(m)}(k, l, s) = \lambda_{k,l}(s) \cdot R^{(m)}(k, l, s) \cdot \prod_{(u,v) \in P_{k,l}^{(m)}, (u,v) \neq (i,j)} a_{u,v}(s) \quad (2)$$

where $R^{(m)}(k, l, s)$ is the probability that a call of class s is not admissible on the first $m-1$ paths given that the m th path is admissible. An algorithm based on two-terminal reliability problem was proposed in [3] to compute $R^{(m)}(k, l, s)$. Two additional approximations of $R^{(m)}(k, l, s)$ was also proposed in [3]:

$$R^{(m)}(k, l, s) = \prod_{r=1}^{m-1} \left(1 - \prod_{(u,v) \in P_{k,l}^{(r)}, (u,v) \neq (i,j)} a_{u,v}(s) \right) \quad (3)$$

with the assumption that admission of a call on a given path is statistically independent of its admission on any other path.

$$R^{(m)}(k, l, s) = \prod_{r=1}^{m-1} \left(1 - \prod_{(u,v) \in P_{k,l}^{(r)} - P_{k,l}^{(m)}, (u,v) \neq (i,j)} a_{u,v}(s) \right) \quad (4)$$

with the assumption that the events $A_{k,l}^{(r)}, r = 1, 2, \dots, m-1$ are statistically independent, where $A_{k,l}^{(r)}$ is the event that $P_{k,l}^{(r)}$ is admissible given $P_{k,l}^{(m)}$ is admissible.

To complete the fixed point approximation, we need a mapping from arrival rates to admissibility probabilities. An approximation first introduced in [6,7] will be used. The probability of admitting a call of class s to link (i, j) is

$$a_{i,j}(s) = 1 - \sum_{n=C_{i,j}-b_s-r_s+1}^{C_{i,j}} E_{i,j}(n) \quad (5)$$

where $E_{i,j}(n)$ is the equilibrium distribution of the following one-dimensional Markov chain:

Letting $\alpha_{i,j}(s)$ denote the average number of calls of type s in progress on link (i, j) ,

$$\alpha_{i,j}(s) = a_{i,j}(s) \cdot \nu_{i,j}(s) / \mu_{i,j}(s) \quad (6)$$

consider the one-dimensional Markov chain, which, for any state n and call class s , jumps to:

- state $n + b_s$ with rate $\lambda_{i,j}(s) I(C_{i,j} - n \geq r_s + b_s)$, where r_s is the trunk reservation parameter, and to
- state $n - b_s$ with rate $\nu_{i,j}(s) n \frac{\alpha_{i,j}(s)}{\sum_t \alpha_{i,j}(t)} I(n \geq b_s)$.

We can obtain the admissibility probabilities $a_{i,j}(s)$ by solving the above small system of fixed point approximation. It is shown in [6,7] that this approximation becomes asymptotically exact in overload conditions. More refined approximations are also discussed in [6,7].

3 PNNI ROUTING

PNNI is the protocol that enables the building of multi-vendor, interoperable ATM networks. PNNI is a hierarchical, dynamic link-state routing protocol for building large-scale ATM-based network. In addition, PNNI defines signaling requests to establish point-to-point and point-to-multipoint connections.

UPPER LEVEL

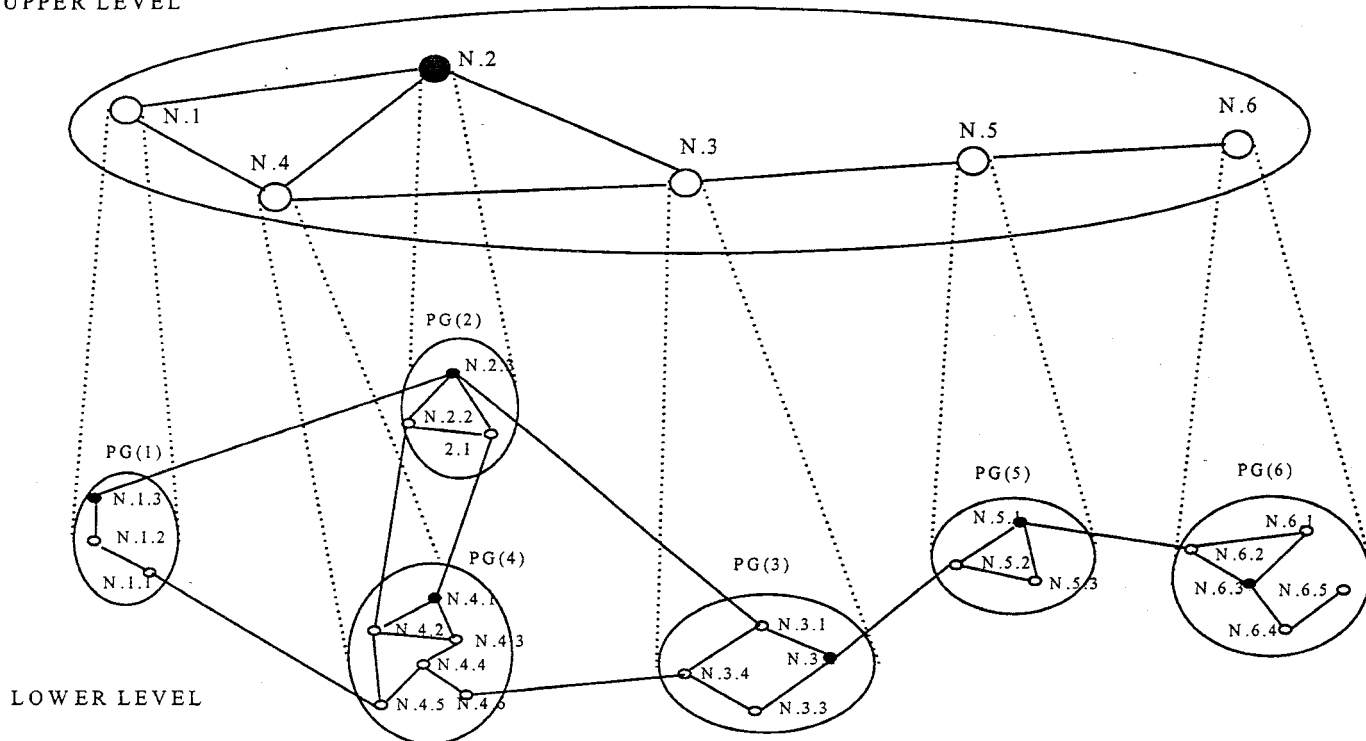


Figure 1. Hierarchical ATM Network with Two Levels

Like OSPF (Open Shortest Path First), PNNI is a hierarchical, link-state routing protocol that organizes switching systems into logical collections called peer group. Neighboring nodes form a peer group by exchanging their peer group identifiers (PGIDs) via Hello packets using a protocol that makes nodes known to each other. If the nodes have the same PGID, they belong to the peer group defined by that particular ID; if their PGIDs are different, they belong to different peer groups. A border node has at least one link that crosses the peer group boundary. Hello protocol exchange occur over logical links (physical link or VPC or SVCC). PNNI defines the creation and distribution of a topology database that describes the elements of the routing domain as seen by a node. This database provides all the information required to compute a route from the node to any address that is reachable in or through that routing domain. Nodes exchange database information using PTSEs (PNNI Topology State Elements). PTSEs contain topology characteristics derived from link or node state parameter information. The state parameter information could be either metrics or attributes. PTSEs are grouped to form PTSP (PNNI Topology State Packet) and PTSPs are flooded throughout the peer group so all nodes in one peer group will have an identical database. Every peer group has a node called peer group leader (PGL). There is at most one active PGL per peer group. The PGL will represent current peer group in the parent peer group as a single node. The PGL will also flood the PTSEs in parent peer group to the current peer group. Apart from its specific role in aggregation and distribution of information for maintaining the PNNI hierarchy,

the PGL does not have any special role in the peer group. Currently, PNNI support 104 hierarchical levels.

Call establishment in PNNI consists of two operations: the selection of a path and the setup of the connection state at each point along that path. PNNI uses source routing for all connection setup requests. The path is encoded as a Designated Transit List (DTL) which is explicitly included in the connection setup request.

The PNNI network can be described by a hierarchical model. The network is represented by a graph $G(V,E)$, where V is the set of layer 1 nodes, and E is the set of links interconnecting them. $G(V,E)$ will be referred as the actual network. Layer 1 nodes are clustered to form layer 2 nodes, which are clustered into layer 3 nodes etc., up to the highest. Say L -th, layer. The PNNI routing hierarchy also allows asymmetries in the sense that a set of layer i nodes can be directly clustered into a layer $i+n$ node with $1 \leq n \leq L-i$, so for a given lower level peer group its parent peer group can simultaneously be a grandparent or great-grandparent peer group to some other lower level peer group. Completion is achieved until the entire network is encompassed in a single highest level peer group.

4 A HIERARCHICAL REDUCED LOAD APPROXIMATION FOR PNNI ROUTING

We now describe a reduced load approximation method for ATM networks with PNNI routing. We assume that the effective bandwidths are used for admission control and routing. PNNI routing, designed for ATM networks, has a hierarchical

structure. Currently, up to 104 levels are allowed, but 3-6 levels are suffice to a large network. For simplicity, we will only consider the case of two level hierarchies. Figure 1 shows a two level hierarchical ATM network. At the lower level, each node represents a switching system. At upper level, each node represents a collection of one or more nodes at lower level (or called a peer group).

A path in PNNI routing is called a hierarchically complete source route. Such a path is not a fully detailed source route because it does not contain the details of the path outside the originator's peer group. Instead, those portions of the path are abstracted as a sequence of logical group nodes to be transited. When the call setup arrives at the entry switching system of a peer group, that switching system is responsible for selecting a source route describing the transit across that peer group.

Consider a path between node N.1.2 and node N.6.4 in Figure 1. PNNI routing algorithm will build following DTLs:

DTL:[N.1.2, N.1.1]
DTL:[N.1,N.2,N.3,N.5,N.6]

which includes nodes (and links) in lower level and logical nodes (and logical links) in upper level. A two step method is needed to estimate the admissibility probabilities and arrival rates on links along this path.

A. Approximation at Upper Level (Step 1)

- A logical link at upper level represents one or more links between the same peer groups at lower level. The bandwidth of one logical link is the sum of the bandwidth of each link it includes.

A logical node at upper level represents a peer group at lower level. In the reduced load approximation model, a physical node does not contribute to the end-to-end blocking probability estimation. Since a logical node may include a set of nodes and links, it will contribute to the end-to-end blocking probability. For the purpose of reduced load approximation, a nodal aggregation mechanism that captures the blocking probability of the logical node is needed.

There are some conventional methods for topology aggregation [8-10]. In the *Symmetric-Node* approach, a given topology with multiple nodes is collapsed into one virtual node. This approach offers the greatest reduction of information, but it does not adequately reflect any connectivity information in the original topology. The *Full-Mesh* approach uses a logical link between each pair of border nodes to construct the aggregated topology. The draw back of this approach is the number of internal links increases as the square of the number of border nodes. A compromise between the two extreme approaches is a *Star* approach. In the *Star* approach, the virtual node, that is the center of the "star", is explicitly connected to the border nodes via logical links which are not necessarily identical. This approach has the complexity of an order that is linear in the number of the border nodes. However, it offers a limited flexibility for reflecting asymmetric topology information.

We use the Complex Node Representation [1] for our purpose of reduced load approximation at upper level hierarchy. The complex node representation provides aggregated topology information not only for traversing the logical node, but also routing to and from the inside of the logical node. The interior reference point of a logical node is referred to as a *nucleus*. A border reference point of a logical node is known as *port*. A logical connectivity between the nucleus and a port is referred to as a *spoke*. Exceptions can be used to represent particular ports whose connectivity to the nucleus is significantly different from the default. Two ports may also be connected via a *bypass*. The complex node representation of logical node N.4 or peer group PG(4) is shown in Figure 2.

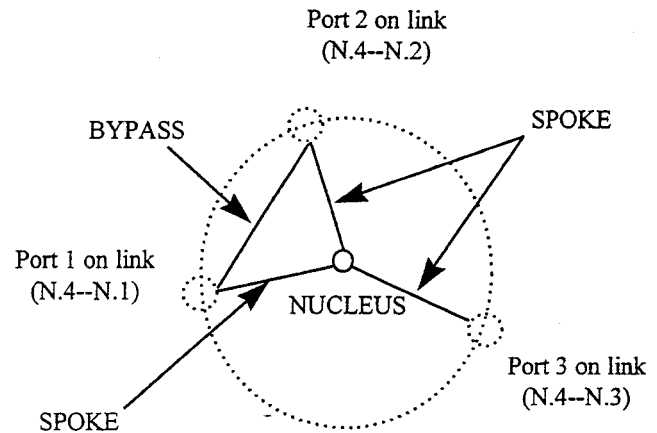


Figure 2. Complex node representation of N.4 in Figure 1

Now replace all the logical node in upper level with their corresponding complex node representations, we get a network with $N^{(u)}$ nodes. A node is either a nucleus or a port of a logical node. A link $(i^{(u)}, j^{(u)})$ is a logical link between peer groups or a spoke or bypass within a complex node representation. Suppose node $i^{(u)}$ of the $N^{(u)}$ nodes represents a set of lower level nodes $G(i^{(u)})$. Then the arrival rate between node pair $(i^{(u)}, j^{(u)})$ is

$$\lambda_{i^{(u)}, j^{(u)}}^{(u)} = \sum_{i \in G(i^{(u)}), j \in G(j^{(u)})} \lambda_{i,j}$$

The computation of link capacity $C_{i^{(u)}, j^{(u)}}^{(u)}$ depends on the type of the link. For logical link between peer groups, the capacity is the summation of the capacity of the lower level links that the logical link represents. For spoke or bypass within complex node representation, the guideline of PNNI topology aggregation [1][9] should be followed. We use a conservative scheme that uses "worst case" parameter, i.e., the smallest

capacity between a nucleus and a port (spoke case) or a port and a port (bypass case). With these parameters, the reduced load approximation algorithm described in section 2 can be used to calculate the admissibility probabilities $a_{i^{(u)},j^{(u)}}^{(u)}$ and aggregated link arrival rates $v_{i^{(u)},j^{(u)}}^{(u)}$ for the upper level network with complex node representations.

B. Refinement at Lower Level (Step 2)

Step 1 gives us the reduced load approximation of the upper level network, so we can compute the blocking probabilities between logical nodes. Now suppose we want to compute the blocking probability between two lower level nodes such as N.1.2 and N.6.4, some refinements are needed. For originating and destination peer group, no aggregation is used, i.e., the logical node N.1 and N.6 will be represented by their corresponding lower level peer groups. For intermediate logical nodes aggregation is used, i.e., complex node representation is used to represent the intermediate logical nodes. The resulting network is shown in Figure 3.

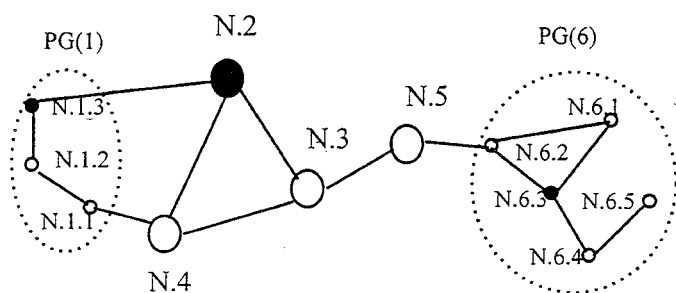


Figure 3. Refinement for Originating and Destination Peer Groups

Now the reduced load approximation algorithm described in section 2 is used to refine the estimation for originating and destination peer groups (PG(1) and PG(6) in Figure 1), i.e., the admissibility probabilities and aggregated link arrival rates for links in originating and destination peer groups are estimated. The admissibility probabilities and aggregated link arrival rates for links other than the originating and destination peer groups will not be estimated and the values we get from step 1 will be used instead. The same refinement can be done for any originating and destination peer group pairs.

This hierarchical scheme follows the PNNI routing framework and results a fast and scaleable reduced load approximation for ATM networks with PNNI routing.

5 CONCLUSION AND FUTURE RESEARCH

The hierarchical approximation scheme proposed in this paper extend the reduced load approximation algorithm to ATM

networks with PNNI routing. Compared to original reduced load approximation algorithm, this scheme uses coarse to fine iteration, so it is fast and scaleable to large networks.

One possible future area of research is to extend the scheme to multiple level of hierarchies. Another interesting problem is to take into account the traffic of topology database flooding and build a more realistic blocking probability model for PNNI routing in ATM networks.

REFERENCES

- [1] ATM Forum, "PNNI Specification 1.0," af-pnni-0055.0000, March 1996.
- [2] F.P. Kelly, "Loss Networks," *The Annals of Applied Probability*, Vol. 1, No. 3, pp. 319-378, Aug. 1991.
- [3] A.G. Greenberg and R. Srikant, "Computational Techniques for Accurate Performance Evaluation of Multirate, Multihop Communication Networks," preprint, 1995.
- [4] S.P. Chung and K.W. Ross, "Reduced Load Approximations for Multirate Loss Networks," *IEEE Trans. on Communications*, Vol. 41, No. 8, pp. 1222-1231, Aug. 1993.
- [5] S.P. Chung, A. Kashper and K.W. Ross, "Computing Approximate Blocking Probabilities for Large Loss Networks with State-Dependent Routing," *IEEE/ACM Trans. On Networking*, Vol. 1, No. 1, pp. 105-115, Feb. 1993.
- [6] N. Bean, R.J. Gibbens and S. Zachary, "Analysis of large single resource loss systems under heavy traffic, with applications to integrated networks," *Adv. Appl. Prob.*, March 1995.
- [7] N. Bean, R.J. Gibbens and S. Zachary, "The Performance of single resource loss systems in multiservice networks," preprints, 1994.
- [8] Whay Lee, "Topology Aggregation in PNNI Routing," ATM Forum, atm94-0874, Sept. 1994.
- [9] Whay Lee, "Some Guidelines for PNNI Topology Aggregation," ATM Forum, atm95-0153, Feb. 1995.
- [10] Whay C. Lee, "Topology Aggregation for Hierarchical Routing in ATM Networks," *ACM SIGCOMM*, pp. 82-92, 1995.
- [11] ARL Advanced Telecommunications /Information Distribution Research Program (ATIRP) Annual Conference Proceedings Vol. I & II, January, 1997.