ISSN No. 0976-5697
Volume 5, No. 7, September-October 2014
International Journal of Advanced Research in Computer Science
CASE STUDY AND REPORT
Available Online at www.ijarcs.info

Load balancing Technique for VOIP in MPLS Networks

AR.Lakshman, KM.Praven
Computer Science and Engineering
R.M.K Engineering College
Chennai, Tamilnadu. India

Abstract— In recent years, there has been a phenomenal growth in the commercial deployment of delay-sensitive services, such as voice and video services, over the Internet Protocol. Regardless of the advances that have been witnessed in networking technologies, effectively utilizing the network resources to accommodate such services remains a major and challenging task for network researchers and designers. In MPLS networks, multiple label switched paths are (LSP) are established between MPLS ingress and egress nodes to enhance the network performance and QoS for subscribers. During unbalanced situation in the network, the flows are routed through multiple paths.

Keywords- MPLS, load balancing, , VoIP and QoS, broadband scenario, Voice, multipath dispersion, , IP network.

I. INTRODUCTION

Voice over IP (VoIP) commonly refers to the communication protocols, technologies, methodologies, and transmission techniques involved in the delivery of voice communications and multimedia sessions Internet Protocol (IP) networks, such as the Internet. VoIP is challenging communications for small and large businesses. Using an internet connection, businesses are able to connect to other offices, call long distance and manipulate incoming phone data in ways that previously would have cost 10 or 20 times these new solutions. Voice over Internet Protocol (VoIP) is often confused with Internet telephony. VoIP is voice transmitted as packets over a data network, whereas Internet telephony refers to voice transmitted as packets over the public Internet—a special case of VoIP. [1]

A general methodology involved in VOIP :

a. Conversion of the caller’s analogue voice signal into a digital format
b. Compression and translation of the digital signal into discrete Internet Protocol packets
c. Transmission of the packets over the Internet or other IP-based network
d. Reverse translation of packets into an analogue voice signal for the call recipient. VoIP has several advantages.

a) Cost reduction
   (a).Toll by pass
   (b).Wan cost reduction
b) Operational improvement
   (a).Common network infrastructure
   (b).Simplification of routing administration
c) Business tool integration
   (a).Voice mail, email and fax mail integration
   (b).Web call
   (c).Mobility using ip
d) New service
   (a).New integrated applications

The quality of packet audio is largely determined by the mouth-to-ear delay and the packet loss. The contribution of this paper is to study and provide techniques to improve the packet audio quality. To achieve our goal of good quality audio communication, some changes might be needed to the present Internet. The degradation of voice quality occur when using a multi-user packet switched network and it is the fundamental problem. Unpredictable short term loads, lack of guarantees on network performance, lack of control over the end systems and stringent requirements on the voice quality make VoIP a challenging application to realize successfully on the Internet. [4]

Qos is defines as “the set of technologies that enables network administrators to manage the effects of congestion on application traffic by using network resources optimally, rather than by continually adding capacity” [5]. The Quality of Service (QoS) research is orthogonal to the investigations being carried out by the network community. These investigations focus on changing the packet switching techniques to be more reliable, more timely and more fair. This is especially the case for time sensitive traffic such as voice. Protocols have been developed to signal routers and end systems that certain data types need to be treated differently, again in the case of voice traffic often at higher priority. We look at allocating resources given the current conditions of the network or adapting to it, also by measuring the current state so that we can make decisions based on these measurements rather than assuming the certain functionality will be available. A media gateway plays a critical role of interoperability between packet networks and the existing telephone networks. Inevitably, it involves many complex processes, such as packetization/depacketization of voice frames, jitter smoothing, error concealment, package loss etc.

Three fundamental concepts affecting real-time data transmission must be considered while designing the IP network for audio and video data. These are network provisioning, queuing, and classifying.

a. ***Provisioning***–provisioning the network simply means installing more network bandwidth or capacity than is actually needed for all of the audio, video, and regular data applications that will run over the network.
b. ***Queuing***–Buffering issues may be overcome by enabling separate voice and video data queues in the network switches and routers. Separate queues allow

time critical data such is audio and video to be transmitted in a priority fashion.

c. ***Classifying***–Several different schemes currently exist for providing priority to network packets. These include Resource Reservation Protocol (RSVP), IP precedence, differentiated services (DiffServ), and Multi-Protocol Label Switching (MPLS).

II. RELATED STUDY

A. *Load Balancing:*

Load balancing is a computer networking methodology to distribute workload across multiple computers or a computer cluster, network links, central processing units, disk drives, or other resources, to achieve optimal resource utilization, maximize throughput, minimize response time, and avoid overload.[1] Using multiple components with load balancing, instead of a single component, may increase reliability through redundancy. The load balancing service is usually provided by dedicated software or hardware, such as amultilayer switch or a Domain Name System server.

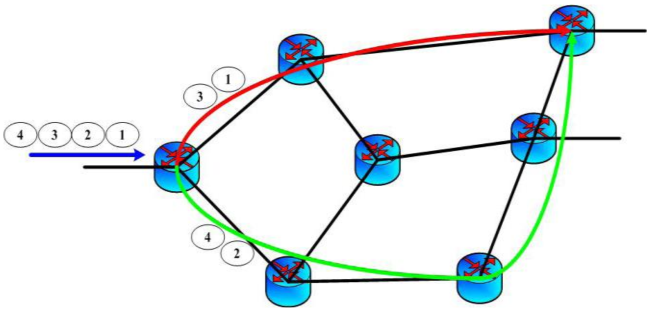


Figure 1. Load Balancing

This section presents a formulation of the load balancing issue and a proposition in a distributed multipath context. Load balancing is common in ISP networks. There exists several theoretical propositions, but only the simplest ones are used in real environments. We tried to find a compromise between computational overhead and reactivity[2]. Indeed, according to the time scale of our measurements to analyze the network activities, it is very difficult to quickly react to fast and strong load oscillations using a complex algorithm.[3]

In this paper we consider load balancing among variable rate flows such as TCP flows. Bandwidth measurement may allow preventing congestions. We choose a load balancing scheme which favors the minimal cost path utilization until significant trouble occurs. Our TE load balancer prevents routers from recalculating unnecessary proportions so often that it could lead to unwanted oscillations. In our context, the objective of real time measurements is to produce a set of proportions corresponding to the quantity of load to share among several next hops for the same destination. When the network is weakly loaded, it is preferable that routers only use one of their best next hops (the minimal cost path) in order to use less resources.[4]This set of proportions is associated to a specific destination and has to verify several conditions.

Formally, we denote, the vector of global (local and transit traffic) proportions according to a destination d on a router s which has n possible next hops to reach dj denotes the rank of the path according to metric C given in table I.[5] We also denote proportions of traffic coming from interface p sent via NHj .

These variable are subject to the following constraints[6]:

$$\sum_{j=1}^{n} x_j^d = 1 \qquad (1)$$

$$\forall p \in I \quad \sum_{j=1}^{n} x_j^d(p) = 1 \qquad (2)$$

$$\forall j \in 1,\ldots n \quad \sum_{p=1}^{k^-(s)} \frac{x_j^d(p) \times V_d(p)}{V_d^T} = x_j^d \qquad (3)$$

Where:

a. x_j^d is the vector of global propotions according to a destination d on a router s which has n possible next hop to reach d.
b. V_d^t denotes Load from p
c. V_d (p) denotes total load aimed at destination d.

Equations (1) and (2) imply the consistency of global and per incoming interface proportions.[7] Equation (3) indicates that the sum of incoming proportions reported to the quantity of traffic they have to support depends on global proportions[8]. We define the function U(l), where l an outgoing link of s, as the total traffic on link l divided by its capacity c_l

$$U(l) = \sum_{p \in I, d \in N}^{\{j|NH_j(s,d)=l.x\}} \frac{x_j^d(p) \times V_d(p)}{c_l}$$

Where

•x_j^d denotes the portion of traffic coming from p which is sent via the link l towards d

A global heuristic to improve network usage is to minimize the maximum link utilization in the network,[9] i.e:

$$\min \max_{l \in L} U(l)$$

The objective of this optimization problem is to anticipate congestion by minimizing the load of highest loaded links. The idea is that when links are weakly loaded, network response time is globally better. Such a formulation implies linear programming in order to optimize this global objective function. [10]

This is unsuitable for a distributed and high performance computation, especially to produce quick local decisions when the traffic is unpredictable. We choose to use a purely local incremental heuristic to approach desired proportions. First, we decide to not consider destination and incoming interface in our measurements, in order to reduce complexity. Each routers only needs to measure the load, denoted ul = U(l) × cl during a chosen time scale t, for each of its outgoing links. The load balancer reacts only when a link l is stressed according to a given threshold: ul > α×cl. Actually, the choice of a time scale is a fundamental issue. For example, a gigabit incoming traffic can fill up a queue of 600 000 bits (75 packets of 1000 bytes) in 0.6 milliseconds.

The monitoring period should be strictly smaller than the millisecond to compute new proportions which allow to dynamically avoid loss. So we aim at preventing only persistent congestions during more than a second. With our load balancing scheme, each router only needs to compute the load of each of its links and to determine which one carries the most critical load.[11]

B. Traffic Engineering with MPLS:

The emergence of MPLS with its efficient support of explicit routing provides basic mechanisms for facilitating traffic engineering. Explicit routing allows a particular packet stream to follow a predetermined path rather than a path computed by hop-by-hop destination-based routing such as OSPF. With destination-based routing as in traditional IP network, explicit routing may be provided by attaching to each packet the network-layer address of each node along the explicit path. This approach generally gains excessive overhead. In MPLS, a path (known as a LSP) is identified by a concatenation of labels which are stored in the nodes. As in traditional virtual-circuit packet switching, a packet is forwarded along the LSP by swapping labels. Thus, support of explicit routing in MPLS does not entail additional packet header overhead.[7]

C. MATE (Measurement Based):

Destination-based forwarding in traditional IP routers has not been able to take full advantage of multiple paths that frequently exist in Internet Service Provider Networks. As a result, the networks may not operate efficiently, especially when the traffic patterns are dynamic. The basic idea of MATE is as follows. The ingress node of each LSP periodically sends probe packets to estimate a congestion measure on the forward LSP from ingress to egress. The congestion measure can be delay, loss rate, or other performance metrics. Each ingress node then routes incoming traffic onto multiple paths to its egress node in a way that equalizes the *marginal* congestion measure (their derivatives). That is, traffic will be shifted from LSPs with higher marginals to LSPs with lower marginals. In equilibrium all LSPs that carry any flow will have minimum and equal marginals. As equalizing the marginal measure minimizes the total congestion measure of the entire MPLS network. Fig. shows a functional block diagram of MATE located at an ingress node.

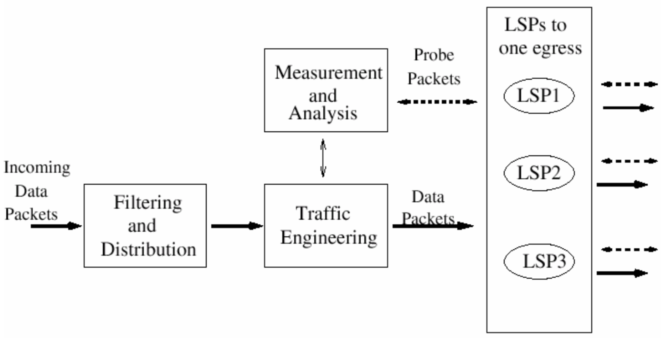


Figure 2. Traffic Engineering

D. Traffic Split ratio with MPLS:

It is known that total traffic throughput in a network, hence the resource utilization, can be maximized if the traffic demand is split over multiple paths. However, the problem formulation and practical algorithms, which calculate the path and the traffic split taking the route considerations or policies into consideration, have not been much touched. This paper proposes practical algorithms that find near optimal paths satisfying the given traffic demand under constraints such as maximum hop count, and preferred or not preferred node/link list. The mixed integer programming formulation also calculates the traffic split ratio for multiple paths. The problems are solved with the split ratio of continuous or discrete values. However, the split ratio solved with discrete values (0.1, 0.2 etc) are more suitable for easy implementation at the network nodes. The proposed algorithms are applied to the MPLS that permits explicit path setup. The paths and split ratio are calculated off-line, and passed to MPLS edge routers for explicit LSP setup. The experiment results show that the proposed algorithms are fast and superior to the conventional shortest path algorithm in terms of maximum link utilization, total traffic volume, and number of required LSPs.

E. Failure Recovery with MPLS:

The proposed multiple QoS path computation algorithm searches for maximally disjoint (i.e., minimally overlapped) multiple paths such that the impact of link/node failures becomes significantly reduced, and the use of multiple paths renders QoS services more robust in unreliable network conditions[8]. The algorithm is not limited to finding fully disjoint paths. It also exploits partially disjoint paths by carefully selecting and retaining common links in order to produce more options. Moreover, it offers the benefits of load balancing in normal operating conditions by deploying appropriate call allocation methods according to traffic characteristics. In all cases, all the computed paths must satisfy given multiple QoS constraints. Simulation experiments with IP Telephony service illustrate the fault tolerance and load balancing features of the proposed scheme.

III. PROPOSED SYSTEM

A. Multipath Routing:

Multipath routing is a routing technique of networks. It uses multiple alternative paths through a network. This provides a variety of benefits such as fault tolerance, increased bandwidth or improved security. Sometimes the multiple paths can be overlapped, edge-disjoined or node disjoined with each other. The implementation of multipath routing deployment is practically very difficult. Much research is needed to overcome these facilities.[9]

a. Divide the message into multiple pieces and routes them to the destination through the selected multiple paths.
b. The dynamic source routing protocol AOMDV is used in Multiplex and Multicast environment
c. Instead of sending through a single path, shares are sending through multiple paths with minimal threshold value.
d. The enemy cannot reconstruct the original message-very difficult to decode
e. Confidentiality and privacy are at greater risk in VoIP systems unless strong controls are implemented and maintained.

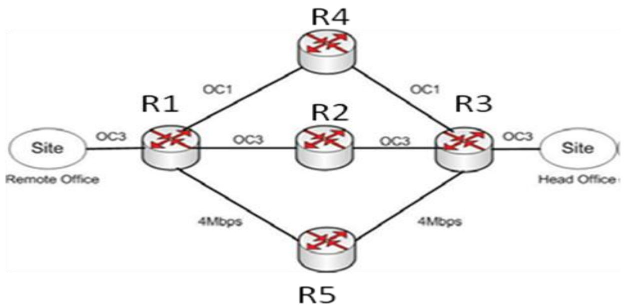


Figure 3. Multipath Routing

Though there are degradation it is still acceptable as delay is within 150ms, this degradation is expected to increase if more data services are added (such as FTP); since the mail server is the only source for data traffic in our case. In the case of multipath routing, although the delay and variance are still in the permissible range of affordable service for utilization up to 80%, but when utilizations of the links are increased, the delay and variance exceeded the threshold value. Also, similar to the single-path routing, when more services are added, the network design has to be revisited again to ensure the expected QoS for voice services. [10]

Table I shows the packet loss results for both cases: single-path and multipath scenarios.

TABLE I. DELAY, JITTER AND PACKET LOSS RESULTS OF SCENARIOS FOR SINGLE PATH AND MULTIPATH WITH UNEQUAL LOAD BALANCING

	Utilization (%)	End-to-End Delay (ms)	Jitter	Loss (%)
Single Path	0	71.425	0.0	0
	50	71.428	0.0	0
	80	71.439	2×10^{-10}	0
	90	71.456	6×10^{-10}	0
	95	71.471	12×10^{-10}	0
Multi Path	0	82.02	5.478×10^{-5}	0
	50	82.84	5.721×10^{-5}	0
	80	84.03	8.476×10^{-5}	0
	90	2430	3.52	1.67
	95	4750	14.6	6.64

In another scenario, one of the links passing through R2 has been failed during the simulation. This is done intentionally to check the degradation in the voice service due to link failure of the link with the highest bandwidth in the backbone. The results are presented in Table II.

TABLE II. DELAY RESULTS OF THIRD SCENARIO FOR MULTIPATH FLOW OF PACKETS WITH FAILURE IN ONE OF THE LINK

	Utilization (%)	End-to-End Delay (ms)	
		before failure	after failure
Multi Path	0	82.30	84.87
	50	82.50	85.09
	80	84.43	95.6
	90	794.71	4602.6
	95	1562.91	7364.23

The Table III is to demonstrate the effect of unequal load balancing when a high bandwidth link fails during operation. A small degradation in the service has been observed which is acceptable as it still ensures the required QoS if the utilization of the other backbone links is less than 80% (similar to the case when no link failure occurs). It should be noted that due to unequal load balancing, packets can be received out of order. In such case, it is the responsibility of the RTP protocol to buffer the received packets and rearrange them before depacketization. It should be noted that in all simulations for the last test, when failure occurs, jitter increased exponentially which is later stabilized and decreased as shown in Table III.

TABLE III. JITTER RESULTS OF THIRD SCENARIO FOR MULT-PATH FLOW OF PACKETS WITH FAILURE IN ONE OF THE LINK

	Utilization (%)	Jitter	
		before failure	after failure
Multi Path	0	5.516×10^{-5}	6.35×10^{-5}
	50	5.69×10^{-5}	6.55×10^{-5}
	80	8.27×10^{-5}	22.26×10^{-5}
	90	0.38721	8.1953
	95	1.65562	22.621

B. Applying Multipath Routing To Mpls Networks:

To achieve load balancing and improve link utilization VoIP packets are dispersed into multiple disjoint paths. For a given source-destination pair multiple set of disjoint paths are computed in an MPLS network [1]. Multipath routing increases fault-tolerance and reliability. MPLS router can split the same label traffic flow into different paths with the given traffic engineering constraint. QoS constraints like minimum delay and maximum bandwidth are considered for splitting a given flow dynamically into these multiple paths. In the proposed multipath routing MPLS ingress router evaluates the flow rate and destination address and inserts a label into each packet. Core routers do not maintain any per flow state information. They make forwarding decision based on their labels. Multipath routing tries to achieve load balancing by splitting the traffic into multiple paths as shown in Figure.

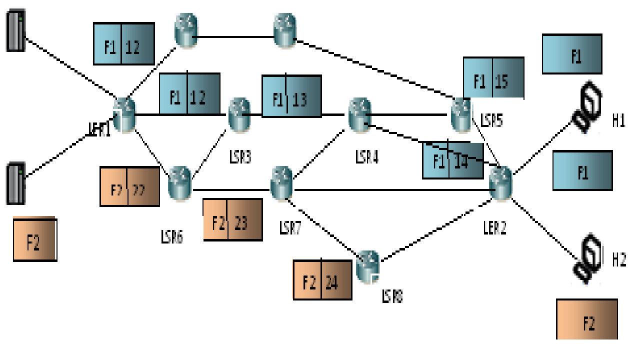


Figure 4. MPLS Network

Ingress router LSR1 receives packets from server1 (flow1) and inserts a label into these packets. LER1 divides the flow1 from server1 into two Label Switched paths (LSP) and distributes the packets concurrently. LSP1 for the flow1 follows LER1—LSR1—LSR2—LSR5—LER2 and LSP2 for the flow1 follows LER1— LSR3—LSR4—LER2 and

LSP1 for flow2 from server2 follows LER1—LSR6—LSR7—LSR8—LER2. Finally LER2 removes the labels and transmits the packets to host1 and host2 [1]. Algorithm for load balancing and multipath routing is shown in Figure

Algorithm (For load balancing and multipath routing)

1. Find P, a set of disjoint loop less path from S to D

2. Find path p_k from the set of disjoint loop less path (P) such that $bw(p_k) = Bw_{req}(f1)$ and $d(p_k) = d_{req}(f1)$

3. Select k backup path from the set P { $p_1, p_2, p_3 \ldots P_k$ } where k is the maximum number of edge disjoint paths

Which statisfies the following bw and delay requirements

$$Bw(p_1) + Bw(p_2) + Bw(p_3) + \ldots\ldots Bw(p_k) = Bw_{req}(f1) \text{ and}$$
$$d(p_1) + d(p_2) + d(p_3) + \ldots\ldots d(p_k) = d_{req}(f1)$$

3. For load balancing

If Queue occupancy >= BOT

{

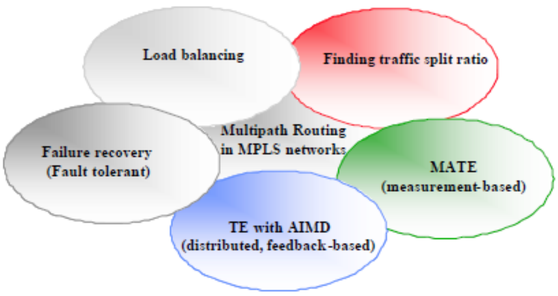
Split Traffic flow among all these k disjoint paths

Disperse the packets using round robin dispersion. }

else

{

Disperse into the a path p_k }



Multipath routing in MPLS networks

IV. CONCLUSION

Multipath routing can be effectively used for maximum utilization of network resources. It gives the node a choice of next hops for the same destination. The various algorithms discussed give solutions for effectively calculating the multipath and ways to minimize delay and increase throughput. Multipath routing is capable of aggregating the resources of multiple paths and reducing the blocking capabilities in QoS oriented networks, allowing data transfer at higher rate when compared to single path. It also increases the reliability of delivery.

V. REFERENCES

[1] Faritha Banu, J. and V. Ramachandran,"Multipath Adaptive Packet Dispersion for Voice Applications",journal of computer science, vol. 8 isuue 4, Page(s) 454-459, DOI : 10.3844/jcssp. 2012.454.459 .2012.

[2] Pascal Mérindol, Jean-Jacques Pansiot, Stéphane Cateloin,"Improving Load Balancing with Multipath Routing" IEEE ,page(s):1-8 ,ISSN:1095-2055, DOI:10.1109/ICCCN.2008.ECP.30 ,December 2008

[3] Parvinder Singh,Deepshikha Verma & Sukhbir Singh,"REVIEW OF VOIP TECHNIQUES",International Journal of Information Technology and Knowledge Management, Vol. 2, No. 2, pp. 397-399,July-December 2009

[4] Hong Li and Lorne Mason,"Multipath routing with adaptive playback scheduling for Voice over IP in Service Overlay Networks",IEE,Page(s):1 5, DOI: 10.1109/ SARNOF.2008. 4520089 2010

[5] Gyu Myoung Lee, Jin Seek Choi,"A survey of multipath routing for traffic engineering", Information and Communications University(ICU)

[6] Sanda Dragos & Radu Dragos,"Bandwidth Management in MPLS Networks" First Joint IEI/IEE Symposium on Telecommunications Systems Research,Dublin, Ireland, 27th November 2001.

[7] Stefano Avallone, Vittorio Manetti, Marina Mariano and Simon Pietro Romano,"A splitting infrastructure for load balancing and security in an MPLS network", Testbeds and Research Infrastructure for the Development of Networks and Communities, 2007. TridentCom2007, IEE Page(s):16, DOI:10.1109/TRIDENTOM.2007.4444681

[8] Mingui Zhang, Bin Liu, and Beichuan Zhang,"Load-Balanced IP Fast Failure Recovery", IPOM '08 Proceedings of the 8th IEEE international workshop on IP Operations and Management ,Page(s) 53 – 65,ISBN:978-3-540-87356-3,2008

[9] G. Murugesan, A.M. Natarajan and C. Venkatesh"Enhanced Variable Splitting Ratio Algorithm for Effective Load Balancing in MPLS Networks",Journal of Computer Science,2008

[10] R.N.Pise, S.A.Kulkarni, and R.V.Pawar,"Packet Forwarding with Multiprotocol Label Switching",vol:1,page:587- December 2005

[11] Jinkuan Wang, Cuirong Wang,"An Adaptive MPLS Multipath Transmission Algorithm",2009 International Conference on Research Challenges in Computer Science, DOI:10.1109/ICRCCS.2009.54,December 2009.