CHAPTER 1

INTRODUCTION

The launch of the American Research Project Agency Network (ARPANET) in 1969 was the beginning of the growth of different networks. The World Wide Web (WWW) now connects millions of computers and the number of hosts continues to grow day by day. Many novel networking technologies have been developed in the Internet to provide Quality of Service (QoS) based applications. Traffic engineering has become crucial within these technologies to meet the increasing demand and variety of customers’ requests. Balancing the link loads among all the links of the network with traffic restoration is one of the most essential tasks of traffic engineering. For this purpose, Multi Protocol Label Switching (MPLS) is increasingly utilized to meet the objectives of traffic engineering. Explicit routing of MPLS facilitates a good balancing of the load by splitting the traffic and diverting it from a congested part of the network to a fairly free path in a well controlled way.

1.1 TRAFFIC ENGINEERING IN INTERNET PROTOCOL NETWORKS

The Internet is a collection of nodes and links established for the purpose of delivering Internet Protocol (IP) datagram, from a source host to a destination host. Internet traffic has a rich variety of characteristics depending on location in the network and at what time scale the traffic is observed. Wide area network and Web traffic possess self similar properties. Self similarity
means that traffic behavior is independent of the time scale when the traffic is observed. If the traffic is bursty on the millisecond level it is bursty at the second level also. In this regard, a sophisticated traffic engineering technique is required to balance the dynamic variations of traffic in the network.

The concept of Traffic Engineering (TE) has been used among town planners and road safety engineers over a long period. They are concerned with flow of vehicles congestion, safety, reliability etc. These concepts apply within packet switched networks to balance the network traffic.

Traffic Engineering involves the study of the current traffic and its usage within the network. It explores the possibilities of utilizing other available paths and links in the network and directs the traffic along multiple paths to maximize throughput. This is achieved by incorporating a combination of extensions to the existing Interior Gateway Protocols (IGPs), traffic monitoring tools, and traffic routing techniques.

1.1.1 Performance Objectives of Traffic Engineering

The performance objectives associated with traffic engineering can be classified as either traffic oriented or resource oriented.

Traffic oriented performance objectives include those aspects that enhance the Quality of Service (QoS) of traffic streams such as,

- Minimization of congestion
- Minimization of packet loss
- Minimization of delay
- Maximization of throughput
- Enforcement of service level agreements

Resource oriented performance objectives include the aspects pertaining to the optimization of resource utilization. Efficient management of network resources is the vehicle for the attainment of resource oriented performance objectives. The above objectives can be achieved by balancing the load of all the links in the network effectively.

1.1.2 Limitations of Current Interior Gateway Protocols on Traffic Engineering

The control capabilities offered by the existing IGPs are not adequate for traffic engineering. This makes it difficult to actualize effective policies to address network performance problems. Indeed, IGPs based on shortest path algorithms contribute significantly to congestion problems within the Internet. Shortest Path First (SPF) algorithms generally optimize the performances based on a simple additive metric. These protocols are topology driven; so bandwidth availability and traffic characteristics are not the factors to be considered in routing decisions. Consequently, congestion occurs frequently when the shortest paths of multiple traffic streams converge on specific links or router interfaces, or when a given traffic stream is routed through a link or router interface whose bandwidth is inadequate. The existing techniques to increase the reliability of traffic delivery within the SPF routing system solve the above constraint to some extent. But they do not increase the amount of traffic the network can handle.
1.2 MULTI PROTOCOL LABEL SWITCHING (MPLS) FOR TRAFFIC ENGINEERING

A popular approach to circumvent the inadequacies of current IGPs and balance the network load to accomplish the traffic engineering objectives is through the use of Multi Protocol Label Switching (MPLS). Anoop Ghanwani et al (1999), Awdeche D (1999) and Jong Moon Chung (2000) have done considerable work on traffic engineering and suggested MPLS to deal with traffic engineering in IP networks. The MPLS extends the design space by enabling arbitrary virtual topologies to be provisioned atop the network's physical topology. This can be done in a fairly straightforward manner.

1.2.1 Advantages of MPLS for Traffic Engineering

For traffic engineering MPLS offers the following advantages:

- Explicit label switched paths which are not constrained by the destination based forwarding paradigm can be easily created by the underlying protocols.
- Label Switched Paths (LSPs) can be potentially and efficiently maintained.
- Traffic trunks can be instantiated and mapped onto LSPs.
- MPLS allows both traffic aggregation and disaggregation whereas classical destination based IP forwarding permits only aggregation.
- It is relatively easy to integrate a "constraint-based routing" framework with MPLS.
- A good implementation of MPLS can offer significantly lower overhead than competing alternatives for Traffic Engineering.
Although these capabilities support traffic engineering, it is not really sufficient and additional augmentations are required to foster the actualization of load balancing policies leading to performance optimization of large operational networks.

1.3 LOAD BALANCING

A central function of traffic engineering is to efficiently balance the load among the resources. It is generally desirable to ensure that subsets of network resources do not become over utilized and congested while other subsets along alternate feasible paths remain underutilized. The effective utilization of network resources will improve all other traffic oriented performances.

1.4 TRAFFIC RESTORATION

In addition to the load balancing, the traffic engineering has to provide protection to the traffic flows when any link failure occurs in the network. Backup or protection LSP can be predefined using disjoint paths through the network. This is because when there is a failure on the primary or working LSP, the data can immediately be switched to the backup path. However, the predefined backup LSPs keep the network resources without carrying traffic until there is a failure. In MPLS, as parallel paths are used, exiting pre-established working paths can be used to mitigate this issue.

1.5 PROPOSED WORK

Traffic engineering is concerned with the performance of operational networks. In general, congestion resulting from inefficient resource allocation can be reduced by adopting load balancing policies. When
congestion is minimized through efficient load balancing, the packet loss and the end-to-end transit delay are decreased and the aggregate throughput is increased. Thereby, the qualities of services experienced by end users are significantly enhanced.

Clearly, load balancing is the essential function of traffic engineering. The capabilities provided for load balancing should be flexible enough so that network administrators can implement other policies which take into account the prevailing cost structure and the utility or revenue model. MPLS is strategically significant for load balancing because it can potentially provide most of the functionality available from the overlay model.

Since the dynamic load balancing is a closed loop control system, there is tradeoff between the number of iterations required to reach the balanced state and magnitude of oscillations that are generated due to rebalancing action at the balanced state. The oscillations will create instability in the system. It is very difficult to provide fast load balancing without oscillations. This is an active research area in the MPLS enabled IP networks.

There are many dynamic load balancing algorithms existing in the MPLS enabled IP networks. However, those algorithms concentrate either on load balancing speed or on magnitude of oscillations. Another limitation of the existing algorithms is that they use path cost as control variable to balance the load. The active path cost measurement offers overlapping information of links and this overlapping information leads to packet loss in the network.

The following algorithms are proposed to overcome the limitations and to improve the performance of the MPLS enabled IP networks.
1. Adaptive Splitting Ratio (ASR) Algorithm
2. Enhanced Variable Splitting Ratio (EVSR) Algorithm
4. Enhanced Variable Splitting Ratio with Oscillation Removal (EVSOR) Algorithm

The proposed algorithms are evaluated numerically as well as using network simulator NS-2.

### 1.6 OVERVIEW OF THE THESIS

The following chapters of this thesis are organized as follows: Chapter two presents the literature review of load balancing and restoration in MPLS networks. In Chapter 3 Adaptive Splitting Ratio (ASR) algorithm is introduced to reduce the end to end delay and to increase the throughput of the MPLS networks when the magnitude of traffic variation is small. Even with large variations in the magnitude of traffic, in order to increase the performance of MPLS network Enhanced Variable Splitting Ratio (EVSR) algorithm is proposed in Chapter 4. In Chapter 5, restoration support is incorporated to the EVSR algorithm to provide protection to the traffic carried on different paths. Enhanced Variable Splitting Ratio with Oscillation Removal (EVSOR) algorithm is developed in Chapter 6 to eliminate oscillations completely. Finally, in Chapter 7 the findings of the thesis and the scope for future research are discussed.