How can Multi Plane Routing be used as a Routing Paradigm in Future IP-based Wireless Access Networks?

A. Jaron, Student Member, IEEE, P. Pangalos, and A.H. Aghvami, Fellow, IEEE

Abstract—We are currently entering a phase where big changes in the network traffic might oblige to rethink the design of current architectures. New technologies, however, being introduced into the market, will provide the ability for such changes. In this paper we focus on the case of multi-path-like routing using the increasingly popular concept of Multi-topology routing. This solution applies to IP-based wireless access networks and provides a solution for optimizing network resources and improving performance.

We first formulate the problem of link weight assignment in multi-topology routing using a linear programming problem (LP). Further, a demonstrator including a GUI was implemented to validate the research and show the benefits of using the multi-plane routing approach.

Index Terms—Multi-Plane Routing, IP wireless access networks, load balancing, network optimization, OSPF

I. INTRODUCTION

The drive towards IP-based networking is an increasingly popular concept in computer communications. Traffic on the Internet is growing exponentially due to the evolution of network devices, services and applications that has reached a phase which imposes to rethink network design, making the case for a clean-slate design of future network paradigms, which are often referred to as “post-IP” solutions. Wide area wireless networks are also experiencing rapid growth in terms of demands originating from end users. Currently, there are many efforts underway to provide data services on wireless access networks. The Internet Protocol (IP) is the dominant internetworking protocol in today’s Internet. OSPF is an link-state adaptive protocol in today’s Internet. OSPF is an adaptive link-state routing protocol which operates directly over IP. Forwarding decisions are exclusively based on the destination address in packets IP headers. Therefore, it does not allow explicit routing on one hand, and explicit splitting of traffic on the other hand. Equal Cost Multi-Path (ECMP) is a feature of OSPF that can achieve load balancing by tuning link weights that is comparable to MPLS networks [7]. However, practically, ECMP solely supports even traffic splitting which does not suffice to achieve a near-optimal performance as in MPLS.

To avoid the problems associated with extra complexity of MPLS, link weight changes and even traffic splitting, authors in [6], [8] proposed a new strategy based on related Multi-Plane Routing techniques, such as Multi-Topology OSPF/IS-IS [9], [10]. Also, Wang. et al. [11] claimed that by partitioning the overall network demand into multiple subsets at the edge of the network so that each of them is delivered through dedicated IP routing planes, near-optimal performance could be achieved. MPR allows the routers within an area to maintain several independent logical planes, with independent set of link weights, and hence independent routing tables for each plane. A plane is a subset of the underlying network (or physical topology). It can overlap with another or share any subset of the underlying network.

In this paper, we developed an algorithm in the form of a Linear Programming (LP) problem using the Shortest Path Formulation (P-SP) for constructing a set of logical routing planes with the use of an offline approach. We then specify, for each source-destination pair, a fixed traffic splitting ratio that alleviates the problems associated with link weight
changes, namely dynamic mapping from flows to routing planes may cause traffic instability. Finally, our experiments based on realistic application models have shown that our approach can lower the maximum link utilization (MLU) and the maximum link delay (MLD) by up to 40% and over 90% respectively.

The remainder of the paper is organized as follows. In Sec. II, we introduce the problem of building up the logical routing planes via a Linear Programming problem formulation. We then describe the traffic splitting ratio adjustment problem using a branching factor. A detailed presentation of the GUI Demonstrator used for the simulations and the performance of our strategy are presented in Sec. III. Finally, Sec. IV concludes the paper.

II. MULTI-PLANE ROUTING FRAMEWORK

In this section, we first describe the principle of Multi-Plane Routing with a simple example. We then formulate the problem of the link weight assignment for constructing the Routing Planes (RPs) as an LP problem. Finally, we develop a fixed traffic splitting ratio adjustment.

A. MPR Principle

MPR allows the routers in an Autonomous System (AS) to maintain several overlaid logical planes, hence our choice for the terminology “plane” instead of “topology”. Each plane is a subset of the physical topology where some of the links have been “removed” or “isolated”, that is, they will not be used for carrying traffic. Therefore, a separate routing table is maintained for each plane allowing the network operator to apply independent routing configurations or add a level of granularity to each plane. Data traffic is then mapped to a specific routing plane, and is routed according to the corresponding Routing Information Base (RIB), or routing table. A more comprehensive description on how to forward an incoming packet into a specific routing plane is shown in [6], [11].

The first question is how the planes should be built so that path diversity is maximized. The ultimate objective of our approach is indeed to maximize path diversity between routing planes so that all the routes between a single source-destination pair are sharing as few links as possible. This objective is the core rule of our strategy all the more so we want to lower the maximum link utilization (MLU) and the maximum link delay (MLD) by up to 40% and over 90% respectively. Finally, our experiments based on realistic application models have shown that our approach can lower the maximum link utilization (MLU) and the maximum link delay (MLD) by up to 40% and over 90% respectively.

To summarize, three properties regulate the construction of a set of routing planes:

1. All planes are connected which means, in each plane, there is a valid route between each gateway (GW) – Access Router (AR) pair. All nodes in between are considered transit routers, they are not traffic sources or sinks.
2. Each link must not be used for forwarding in at least one routing plane.
3. Each link must be used in at least one plane. This ensures path diversity is maximized.

It has been found in many operational networks that the network utilization is skewed, i.e., some links are highly utilized while others are not. Since we know that high-utilization leads to an increase in delay, which in our case is even more significant as we used the M/M/1 queueing system, this is not necessarily desirable if a highly utilized link is on the shortest path to the destination.

B. Link Weight Assignment Problem Formulation

In OSPF-based networks, shortest paths (SPs) are computed based on the sum of the link weights along the path between a source and a destination node. Therefore, the link weight setting controls the computation and calculation of the SPs in the network.

Let consider a network that is defined as an undirected graph \( G = (N,A) \), where \( N \) is the node set and \( A \) is the arc (or link) set with capacity \( c(a) \) for each \( a \in A \). \( N_r \subseteq N \) is the access router set which contains all access routers in the network so that each AR \( r \in N_r \). We define the gateway by \( \in N \setminus N_r \). The demand for each pair \((i,j) \in N \times N \) tells the demand \( D(i,j) \) in traffic flow between \( i \) and \( j \). More precisely, in our case, the traffic demand concerns only the demand between the gateway and the ARs. That is, \( D(i,j) = D(g,r) = d, g \in N \setminus N_r, \forall r \in N_r \). And for each demand \( d \in D \), let \( b_d, s_a(=g) \), \( t_d(=r) \), \( \forall r \in N_r \) be the demand bandwidth (size), the source node and the destination node respectively.

As a preparation step, we need to borrow some basic results from linear programming. The details can be found in [12], [13]. Since shortest path is the most elementary routing paradigm and is used as a building block in our work, we would like to first give a brief insight on the shortest path problem in the context of linear programming. For this purpose, let \( \{w_{ij} : (i,j) \in A \} \) be a given set of link weights. We need to define the following variables. For each link \((i,j) \in A \) and for each demand \( d \in D \), let \( \psi_{ij}^d \) represent the portion of demand \( d \) which flows across link \((i,j) \). Then the shortest path problem can be formulated as

Shortest Path Formulation (P-SP)

\[
\begin{align*}
\min & \sum_{d \in D} \sum_{(i,j) \in A} w_{ij} \psi_{ij}^d \\
\text{s.t} & \\
& \sum_{d \in D} \psi_{ij}^d - \sum_{d \in D} \psi_{ji}^d = 0, \\
& \sum_{d \in D} \psi_{ij}^d \leq s_a, \\
& \psi_{ij}^d \geq 0.
\end{align*}
\]

![Fig.1. Example of a five routing plane configuration. Numbers indicate link IDs (left) and link weights for one RP (right).](image)
∀d ∈ D, ∀i ≠ s_d, t_d

\[ \sum_{j:(i,j) \in A} \psi_{ij}^d - \sum_{j:(j,i) \in A} \psi_{ji}^d = D(g, r), \]  
(2)

∀d ∈ D, i = s_d

\[ \sum_{j:(i,j) \in A} \psi_{ij}^d - \sum_{j:(j,i) \in A} \psi_{ji}^d = 0, \]  
(3)

∀d ∈ D, i = t_d

0 ≤ \psi_{ij}^d ≤ 1, \forall (i,j) ∈ A, ∀d ∈ D  
(4)

The objective function (1) is to minimize the total weight of links used. Constraint (2) and (3) are flow conservation constraints. Equation (2) indicates that the traffic flowing into a node must equal the traffic flowing out of the node for any node other than the source node and the destination node for each demand d. Equation (3) signals that the net flow out of the source node is 0, which is the total required bandwidth after scaled by \( b_d \). The last constraint simply restricts all the variables to be non-negative real numbers between 0 and 1.

The routing of demand is solely determined by the SPs which in turn are determined by the weights we assign to the arcs. The problem of Link Weight Assignment in MPR (LWA-MPR) is described as follows. First, let \( K \) be the set of routing planes so that \( K : k = 1, ..., K \). Then let \( \psi_{ij}^d(k) \) be the percentage of traffic for demand \( D(g, r) \) between gateway \( g \in N \setminus N_r \) and access router \( r \in N_r \) flowing through arc \((i,j) \in A \) on routing plane \( k \). Let \( f_{ij}^d \) be the total portion of traffic flowing through arc \((i,j) \in A \) for demand \( d \). The load on arc \( a = (i,j) \in A \) is denoted by \( l(a) = \sum_{d \in \Delta} f_{ij}^d \). Therefore, the utilization of arc \( a \) is \( u(a) = l(a)/c(a) \). Link weight \( \omega_{ij}(k) \) denotes the weight of link \((i,j) \in A \) in routing plane \( k \). We use \( \delta_{ij}(k) \), a binary variable, which is equal to 1 if arc \((i,j) \in A \) is on the shortest path from node \( i \) to access router \( r \) on routing plane \( k \) and is equal to 0 otherwise. The shortest distance from node \( n \in N \) to access router \( r \) on RP \( k \) is designated by \( \gamma_{nr}(k) \). Finally, let \( F \) be the optimal cost and \( r \) the maximum link utilization. LWA-MPR can be written as follows:

**Minimize** \( F = r \)

subject to

\[ \sum_{k \in K} \sum_{(i,j) \in A} \psi_{ij}^d(k) = 1, \]  
(5)

\[ \sum_{j:(i,j) \in A} \psi_{ij}^d(k) - \sum_{j:(j,i) \in A} \psi_{ji}^d(k) = 0, \]  
(6)
\( \forall d \in D, \forall i \neq s_d, t_d, \forall k \in K \)
\[ \sum_{j \in \{i,j\} \in A} \psi_{ij}^d(k) - \sum_{j \in \{i,j\} \in A} \psi_{ji}^d(k) = D(g, r), \]  
(7)
\( \forall d \in D, i = s_d, \forall k \in K \)
\[ \sum_{j \in \{i,j\} \in A} \psi_{ij}^d(k) - \sum_{j \in \{i,j\} \in A} \psi_{ji}^d(k) = -D(g, r), \]  
(8)
\[ \psi_{ij}^d(k) = f_{ij}^d, \; \forall (i, j) \in A, \forall d \in D \]  
(9)
\[ \gamma_{ir}(k) + w_{ij}(k) - \gamma_{jr}(k) \leq \left(1 - \delta_{ij}^d(k)\right) \cdot C \]  
(10)
\[ \forall (i, j) \in A, \forall k \in K, \forall r \in N_r \]
\[ \left(\sum_{d \in E} f_{ij}^d \cdot b_{ij}^d\right) / c(a) \leq r, \; \forall (i, j) \in A \]  
(11)
\[ 0 \leq \psi_{ij}^d \leq 1, \; \forall (i, j) \in A, \forall d \in D \]  
(12)

We try to minimize the objective function which gives us the maximum link utilization \( r \) which is a common metric that is used for load balancing purposes. Constraints (5), (6), (7) and (8) are the same as the P-SP formulation. Equation (9) ensures that the sum of traffic portions for demand \( d \) for all routing planes equals the total demand \( d \). Constraint (10) forces \( \delta_{ij}^d(k) \) to be equal to 1 if arc \( (i,j) \) belongs to the shortest paths to \( r \) in the \( k \)th RP; \( C \) is a large fixed constant. Finally constraint (11) ensures that the utilization of all the links in the network is less than \( r \). Note that to isolate a link from a plane, we simply set its weight to:

\[ I > \sum_{(i,j) \in A} w_{ij}(k) \]  
(13)

This ensures that the link will not be selected for routing traffic. The SP calculation will therefore ignore the link.

III. GUI DEMONSTRATOR AND PERFORMANCE EVALUATION

A. Overview

The demonstrator shown in Fig. 2 depicts the Graphical User Interface (GUI) produced with MATLAB. The simulations are performed throughout a day, namely 24 hours. At each hour of the day, the results are the average of the past hour. The user is asked to follow the order indicated in the GUI. First, the user has to display the network partial visualization. After running the simulation, two graphs are automatically displayed. First, graph number 1 (Fig. 2) shows the total traffic demand at the gateway interface. Second, graph number 2 (Fig. 2) depicts the utilizations of all the links in the network at each hour of the day. The user may then choose what results to plot (panel 4) in order to compare and observe the overall performance. Firstly, on the left-hand side of the demonstrator, one may want to display all the routing planes in new windows. The routing plane topologies will be included in the full version of the paper. “Link Utilization Comparison” and “Link Delay Comparison” buttons present a table comparing the maximum, minimum, mean and standard deviation of link utilizations and delays in the network respectively, based on the chosen routing paradigm (see Fig. 4). “3D Visualization” allows the user to display graphs number 3 and 4 in three dimensions to ease readability. Further, “Max. Link Utilization” presents the evolution of the maximum link utilization against the different routing schemes (the graph will be shown in the full version of the paper). MPR-# denotes Multi-Plane Routing where # represents the number of routing planes. “Max. Link Utilization Evo.” plots the maximum link utilization for each of the routing paradigm against the total network demand normalized by the total network capacity. Finally, the user can plot the link utilization and link delay for one desired link in the network that he judges worth observing (graph number 3 and 4, Fig. 2).

B. Performance evaluation

For comparison purposes, we plotted the link utilizations with a link weight setting proportional to the inverse of the link capacity. This method has been proposed by Cisco and performs well in backbone networks. We also simulated the results with the standard OSPF configuration where all the link weights are set to the same value, for instance 1. This is named “Hop-Count” OSPF. However, due to the specific network topology used in this paper, we observe that...
Cisco’s link weight setting performs as poorly as the hop-count shortest path routing. In Fig. 4, the gain is calculated between the “hop-count” link utilizations and MPR-5. This is the maximum gain achievable with the current simulations. Fig. 4 shows that the mechanism outperforms standard OSPF hop-count shortest paths routing by a considerable margin. Not only the maximum link utilization is reduced from 97% to 40%, but also the minimum, mean and standard deviation of link utilizations are also reduced which confirms that on one hand the load is well balanced across the network and on the other hand the lessened standard deviation indicates that the utilizations of the links tend to be closer to the mean.

The same observation can also be noticed as far as the link delays are concerned. Also from Fig. 4, it is proved that a moderate number of routing planes suffices to reach a good overall network performance as when adding more than three routing planes, the link utilizations decrease slowly and flatten out. This effect is also depicted in Fig. 5 where we see a large gap between MPR-2 and MPR-3.

**IV. CONCLUSION**

In this work we introduced Multi-Plane Routing as a routing paradigm in all-IP wireless access networks. We argue that next generation access networks will be exclusively IP-based, and towards this end, the need for a resilient, efficient and easily implementable routing mechanism is urging. The main advantage of MPR over link weight tuning solutions is the ability to change the routing in response to traffic dynamics without triggering a link-state re-convergence. With MPR, we showed that the maximum link utilization and maximum link delay could be reduced by up to 40% and 90% respectively. We also proved that the overall network performance could be increased with only a reasonable number of routing planes.

We need to note that this is in a nascent stage of research, and more features will be added to the present demonstrator to allow full flexibility for the user on one hand, and mobility patterns and user behavior will be taken into account to strengthen the present study on the other hand.

**REFERENCES**