

A Segment List Selection Algorithm Based on Delay*

*Note: Sub-titles are not captured in Xplore and should not be used

1 st Given Name Surname	2 nd Given Name Surname	3 rd Given Name Surname
dept. name of organization (of Aff.)	dept. name of organization (of Aff.)	dept. name of organization (of Aff.)
name of organization (of Aff.)	name of organization (of Aff.)	name of organization (of Aff.)
City, Country	City, Country	City, Country
email address or ORCID	email address or ORCID	email address or ORCID

Abstract—Segment routing (SR) sets segment routing labels at the entry node of the SR domain to guide the traffic in the network to match the forwarding path with message granularity. The resulting segment routing-based traffic engineering (SRTE) can efficiently and simply perform traffic analysis. Adjustment, because segment routing converts forwarding instructions into label stacks and encapsulates them in segment routing headers, so the choice of SR label stack will greatly affect the effect of SRTE. In this article, we propose a segment routing label stack generation algorithm for the delay parameters in traffic engineering, which can meet the delay requirements of request messages as much as possible. Our algorithm considers node centrality and load balancing, and minimizes the number of segment routing labels to reduce the proportion of message headers.

Index Terms—component, formatting, style, styling, insert

I. Introduction

In the SDN network scenario, the controller, as the management and control plane of the network device, can obtain global network topology information, but all network traffic is regulated by the SDN mechanism, which often lacks some flexibility and timeliness. For example, all network traffic engineering uses SDN-based As a current traffic engineering scheduling scheme commonly used in the industry, MPLS is originally intended to use labels to replace switch IP for fast forwarding. The operations of matching labels and assigning ports can be completed without sending the CPU, so it reduces The switch has a delay in processing and checking the table of messages, which improves the forwarding efficiency. However, the distribution and notification of labels requires internal gateway protocol (IGP) or open shortest path priority protocol (OSPF) for transmission. In addition, when MPLS is working, each next hop is determined, that is, a complete tunnel in the network needs to be reserved for traffic under the MPLS strategy, which needs to be used as a resource reservation protocol for traffic engineering (RSVP) The extended RSVP-TE reserves resources on the IP network. The application program running on the IP network terminal can indicate to other nodes

the nature of the data packet flow (such as bandwidth, jitter, maximum burst, etc.) that the terminal will receive based on RSVP, so that the edge computing MPLS path device can perform the traffic forwarding path according to the demand distribution. However, the current MPLS deployment method has some drawbacks: First, the operating overhead of RSVP-TE as a resource reservation mechanism is usually very high; second, the IP network can originally use multi-path equal-cost routing (ECMP) to achieve native load balancing. However, MPLS, as an explicit source routing, encapsulates all the labels of the routing path in the message header, making it impossible to perform the native multi-path load balancing of the IP network between each hop.

Segment routing provides a new traffic scheduling idea. Unlike source routing, where the path of each hop is planned, segment routing only plans a destination label in the range of several hops, so that the label stack of the segment routing packet header is only used very little. The label can guide the network traffic forwarding. This can be seen as the SDN controller programming the network traffic processing mode. The SR label stack is equivalent to the programmed code, thereby achieving the network programming purpose of operating network data packets.

However, which label needs to be allocated by the segment routing label stack needs to be determined by the control plane with a global perspective. It can be configured through the experience of the operator, or it can be calculated and allocated by the controller in the SDN scenario. At present, the commonly used label stack calculation methods are all based on bandwidth, and in a large environment where the business is more and more sensitive to delay, it is necessary to consider the use of delay as an element to guide label generation, that is, the network layer represented by SR has the effect of delay Provide a certain degree of protection. Therefore, this article will discuss the segment routing label stack generation algorithm that has the best effect of guaranteeing traffic delay based on the global perspective of the SDN controller.

Identify applicable funding agency here. If none, delete this.

II. Motivation

Several papers focus on selecting nodes SR list, but the weight of link they care about is bandwidth. [example here] Bandwidth is easy to obtain in the network, but it cannot reduce latency directly. Congestion control algorithm like Swift has verified that delay is simple and effective for congestion control in the datacenter. So, we are aimed at proposing a routing algorithm based on latency and used it in traffic engineering.

There are several sources of delay in the network, propagation delay, transmission delay, processing delay, and queuing delay. Propagation delay is definite for the time that light travels in the fiber are almost fixed. The farther the distance between the source and destination, the more time it will take to propagate. But the delay caused by the processing of the packet by the switch is almost unpredictable. The more intermediate routers along the way where the packet goes through, the higher the transmission, processing, and queuing delays for each packet. Finally, the higher a load of traffic along the path, the higher the likelihood of the packet being queued and delayed inside one or more switch buffers. In summary, the sum of the delays of processing packets in the switches at both ends of a link is an important basis for measuring the delay of a link. In particular, in-band network telemetry (INT) can measure per-hop sojourn times to provide a more accurate breakdown of delay.

Therefore, we use INT to measure the delay of every link in the network, but the INT is not the focus of this paper. We will use the delay provided by INT to propose a segment routing trans-point selection algorithm.

III. Algorithm

The algorithm for calculating segment routing trans-points is divided into two steps. The first step is the calculation of static core nodes. This step will be carried out regardless of the traffic, and only mathematical calculations are carried out for the topology of the network and the attributes of nodes. The second step is the selection of dynamic nodes. In this step, the network nodes will be weighed and judged based on the delay collected by INT, and finally, draw a conclusion of whether to add a device to the SR segment list.

A. Computing nodes' centrality

Degree centrality for a node v is the fraction of nodes connected to it. For unipartite networks, the degree centrality values are normalized by dividing by the maximum possible degree (which is $n-1$ where n is the number of nodes in G). In the bipartite case, the maximum possible degree of a node in a bipartite node set is the number of nodes in the opposite node set 1. The degree centrality for a node v in the bipartite sets U with n nodes and V with m nodes is

$$d_v = \frac{\deg(v)}{m}, \text{ for } v \in U, \quad (1)$$

$$d_v = \frac{\deg(v)}{n}, \text{ for } v \in V, \quad (2)$$

where $\deg(v)$ is the degree of node v [1]. Betweenness centrality is another kind of centrality. Betweenness centrality of a node v is the sum of the fraction of all-pairs shortest paths that pass-through v . Values of betweenness are normalized by the maximum possible value which for bipartite graphs is limited by the relative size of the two node sets 1. Let n be the number of nodes in the node set U and m be the number of nodes in the node set V , then nodes in U are normalized by dividing by

$$\frac{1}{2}[n^2(s+1)^2 + m(s+1)(2t-s-1) - t(2s-t+3)], \quad (3)$$

where

$$s = (n-1) \div m, \quad t = (n-1) \bmod m, \quad (4)$$

and nodes in V are normalized by dividing by

$$\frac{1}{2}[m^2(p+1)^2 + n(p+1)(2r-p-1) - r(2p-r+3)], \quad (5)$$

where

$$p = (m-1) \div n, \quad r = (m-1) \bmod n. \quad (6)$$

In order to determine which centrality measurement method is more suitable for the scenario of selecting trans-points in this paper, the experiment will verify these two calculation methods respectively.

B. Min-max hop algorithm

On the basis of a set of nodes calculated by the above algorithm, consider the actual situation of the link delay associated with each node in the network. First, take the average of the link delays around each node in the candidate set, and use this value to reflect the current popularity of the node: the higher the average value, the larger the probability that the path to this node will encounter queuing, the lower the average value, the link resources around the node are relatively sufficient.

Since the average delay of the links around the node can be updated as the real-time attribute of the node, the time complexity to obtain this value is $O(1)$, and then the nodes are sorted according to the average experiment of the direct links of the node, the time complexity of the common quicksort is $O(n \log n)$.

This problem needs to be calculated at each time point where the SRv6 segment list needs to be reselected. The mathematical model is described as follows: each node in the network has an undirected edge with delay as the weight. The candidate trans-point is in a set of core node calculated by the last step. Every candidate trans-point cluster the remaining nodes, and finally use a regional delay to represent the delay attribute of a larger range of equipment groups so that the network can be dimensionality reduction of nodes matrices in, it is actually a network node segmentation algorithm. Based on the results of such segmentation, the delay impact factor of

each area can be calculated, and then one or two of the smallest impact factors on the delay can be selected to be organized as the segment list. The algorithm used here is as follows:

In the dynamic scenario described above, we describe such a path selection algorithm (the reason why it is called a path is because in this part, the waypoint set can be regarded as a dimensionality reduction network topology, which is equivalent to this node Select the path with the least delay among fewer network topologies) This algorithm allows us to pre-calculate all possible QoS paths for a given network topology and link delay index, and has a relatively low computational complexity. Specifically, this algorithm allows us to pre-calculate the minimum hop path with the minimum delay for any destination, and its computational complexity can be comparable to the standard shortest path algorithm[3]. The path selection algorithm is based on the Bellman-Ford (BF) shortest path algorithm, which is suitable for calculating the path with the shortest delay for all hops. A characteristic of the BF algorithm is that in its h -th iteration, it identifies the best path (in our context: the minimum delay) among the paths with the most h hops between the source and each target. Therefore, we also take advantage of the fact that the BF algorithm handles the process by increasing the number of hops, thereby fundamentally obtaining the number of hops of the path as the second optimization criterion.

C. Depth of Segment List

Depth of segment list in packet header is a matter of concern. A deep segment list can regulate traffic more accurately, but it will result in a lower effective load rate. Especially in SRv6, where IPv6 is used as a segment, the damage caused by the long segment list will be more serious.

Research has also shown that using 3- or 4SR did not appear to improve the results for ISP topology and was significantly harder regarding computation time and memory requirements. And the results for 3SR and 4SR are not better than those for 2SR, which coincides with the claim about 2SR being near-optimal[2].

Due to the above reasons, our algorithm will only select 2SR nodes in network as the segment list in SRv6 header.

D. Inhibition for link-delay feedback

The SR policy, which reflected in segment list, should not be updated frequently. SR policy in current industrial production is generally modified only when the link fails, which causes some links to have a high load rate, and the packet loss rate increases at the same time.

It is easy to understand that when a node is selected into the segment list due to the above algorithm, the traffic reaching the node will increase, and the delay of nearby links will grow. After the link delay information collected by the INT process, the algorithm will not recommend

placing this node in the segment list. If the SR policy is going to adjust at this time, it will not be conducive to the transmission of the entire network traffic. Therefore, it is necessary to avoid adjusting the flow rate easily, which may cause confusion on a larger scale.

IV. Simulation

In the experiment, we chose bmv2 which can run p4 program as the software switch in the simulation verification. This is because in order to collect the delay of every direct link of the switch, the collection mode of INT needs to be customized, and the software switch of p4 can be better work with it. We have also implemented the forwarding capability in accordance with the SRv6 standard in bmv2. For the control plane, we wrote a python script with thrift-based API as an instance of the controller to complete the calculation and delivery of the segment list.

A. L^AT_EX-Specific Advice

Please use “soft” (e.g., `\eqref{Eq}`) cross references instead of “hard” references (e.g., (1)). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don’t use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in L^AT_EX will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you’ve discovered a new method of counting.

BibT_EX does not work by magic. It doesn’t get the bibliographic data from thin air but from .bib files. If you use BibT_EX to produce a bibliography you must send the .bib files.

L^AT_EX can’t read your mind. If you assign the same label to a subsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

L^AT_EX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it’s supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won’t be any anyway) and it might stop a wanted equation number in the surrounding equation.

B. Some Common Mistakes

- The word “data” is plural, not singular.

- The subscript for the permeability of vacuum μ_0 , and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an “inset”, not an “insert”. The word *alternately* is preferred to the word “alternately” (unless you really mean something that alternates).
- Do not use the word “essentially” to mean “approximately” or “effectively”.
- In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
- Do not confuse “imply” and “infer”.
- The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the “et” in the Latin abbreviation “et al.”.
- The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [?].

C. Authors and Affiliations

The class file is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

D. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use

“figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

E. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. ??”, even at the beginning of a sentence.

TABLE I
Table Type Styles

Table Head	Table Column Head		
	Table column subhead	Subhead	Subhead
copy	More table copy ^a		

^aSample of a Table footnote.



Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

Acknowledgment

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks ...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

References

Please number citations consecutively within brackets [?]. The sentence punctuation follows the bracket [?]. Refer simply to the reference number, as in [?]¹—do not use “Ref. [?]” or “reference [?]” except at the beginning of a sentence: “Reference [?] was the first . . .”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [?]. Papers that have been accepted for publication should be cited as “in press” [?]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [?].

References

- [1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
- [7] M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.