# On the Fast-Convergence of Delay-Based Load Balancing over Multipaths for Dynamic Traffic Environments

---



Citation:

Meng Li, Hiroki Nishiyama, Nei Kato, Kimihiro Mizutani, Osamu Akashi, and Atsushi Takahara,"On the Fast-Convergence of Delay-Based Load Balancing over Multipaths for Dynamic Traffic Environments," 2013 International Conference on Wireless Communications and Signal Processing (WCSP 2013), Hangzhou, China, to appear.

# On the Fast-Convergence of Delay-Based Load Balancing over Multipaths for Dynamic Traffic Environments

Meng Li†, Hiroki Nishiyama†, Nei Kato†, Kimihiro Mizutani‡, Osamu Akashi‡, Atsushi Takahara‡
† Graduate School of Information Sciences, Tohoku University, 980-8579, Japan.
‡NTT Network Innovation Laboratories, 1-1 Hikarinooka, Yokosuka, Kanagawa, 239-0847, Japan.
E-mail:{mengli,bigtree,kato}@it.ecei.tohoku.ac.jp, {mizutani.kimihiro,akashi.osamu,takahara.atsushi}@lab.ntt.co.jp

***Abstract*—Multipath routing enables the source exploit multiple available paths to transfer data to destination. This technique has drawn much attention by efficiently utilizing the bandwidths, preserving packets order and so on. However, these load balancing schemes are not for the delay-related issue and thus unsuited for the real-time applications. To deal with the delay-sensitive features, a load balancing scheme named Effective Delay-Controlled Load Distribution (E-DCLD) has been proposed to lower the end-to-end delay and the associating packet reordering possibility. Nevertheless, to compute the optimal load for each path, this scheme uses gradually approaching method that needs extra convergence rounds, and performs unsatisfactory especially when path status is unstable. In this paper, we propose a Convex optimization-Based Method (CBM) to effectively figure out the best load ratio for each path based on the model of E-DCLD. The proposed method could count out the result at once and overcome the low convergence rate problem of the original solution. Experimental results demonstrate that our solution could significantly decrease the end-to-end packet delay and total packet delay.**
***Keywords*—Multipath transmission; Load balancing; Convex optimization method; Time-varying path status; Queuing theory**

## I. Introduction

Although the emergence of telecommunication techniques provision higher bandwidth for transmission, such resources are still scarce, which gradually become the bottleneck of network and will continue to do so in the years to come. One possible alternative is to make use of multiple logical paths between the ingress and egress gateway so as to improve network performance. The network protocols provide the mapping techniques for such usage. The Internet Protocol (IP) uses Routing Information Protocol (RIP), Enhanced Interior Gateway Routing Protocol (EIGRP) or Open Shortest Path First (OSPF) to establish these paths. In Multi-Protocol Label Switching (MPLS), these paths are called Label Switching Path (LSPs) [1]. In addition to the protocol, the ingress gateway also needs to implement the load adaptor to make decision on how to distribute the load.

In order to fully exploit the transmission ability of these paths, some schemes have been proposed. The majority of them focus on the bandwidth utilization and packet order preservation. However, these methods are not the directly solution for the real-time applications, and cannot guarantee the low delay and ordered packets. The E-DCLD [2] formulate the delay-aimed problem model based on the queuing theory, to strike the lower delay and packet reordering ratio. It uses iterative method to figure the load ratio for each path. Although E-DCLD provides a solution for the real-time transmission, its original solution needs several rounds to convergence to the optimal load distribution ratio. The number of rounds rest with the number of available path. Given that the bandwidth is not stable because of the background traffic and the link failure, it is very likely that the mechanism would be under the convergence condition most of time and the result of such method is not desirable.

This paper shows one can get the optimal load ratio for these reserved paths in one shot. We introduce the Convex optimization-Based Method (CBM), a new load distribution algorithm, to efficiently allocate the load to the available paths. We first prove that such load distribution problem model can be roughly regarded as a convex optimization problem, and then we use the proposed method to solve this problem. Because this method could instantly get the result, it avoids the possible misguidance that the original solution may conduct in its initial stage. It should be noted that our major work is to decide the load distribution ratio. Although the routing decision (what kind of paths between the ingress and egress node will be used) and the rate calculation (how to sample the traffic) also play important roles in traffic engineering, these issues are beyond the scope of this paper. The numerical results demonstrate that the proposed scheme outperforms the original one especially when the path status is time-variant, and the number of paths is relative large. These results are in line with our analysis.

The rest of this paper is organized as follows. Section II introduces the existing load balancing models briefly. In section III, we present the model of the problem to be solved. In section IV, we prove the convexity of the objective function, and use convex optimization method to solve the problem model. The numerical analysis is presented in section V. Finally, section VI summarizes the contribution of this paper and outlines directions for future work.

## II. Related Work

A large body of literatures has been devoted on how to design an effective and efficient load distribution scheme on the ingress gateway in the network. The existing load distribution models implement different strategies in the multipath forwarding mechanism. These models can be coarsely categorized into the following three groups: Flow-based schemes, Packet-based schemes and Hybrid schemes (sub-flow or others).

The Flow-based schemes [3] [4], on one hand, treat flow

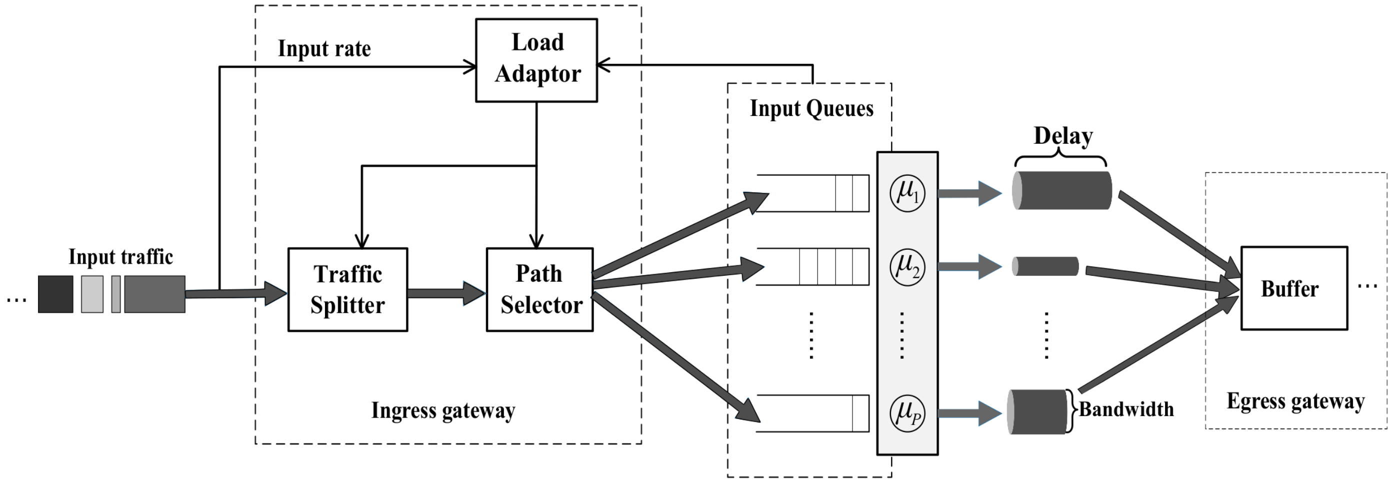


Fig. 1. Functional components of multipath forwarding system.

as the allocating unit. The packets from the same flow will choose the same transmission path. Fast switching (FS) [3] is a widely used flow-based scheme that induces a round-robin path selector. FS need an extra flow table to record the transmission path for each flow. In Direct Hashing (DH) [4] or Table-based Hashing (TH) [5] [6] schemes, the path that a certain flow will choose is determined by the hash value of the packet identifier of this flow. These flow-based schemes have the advantage of preserving the order of packets. However, their average end-to-end delays are often undesirable. The rationale is twofold: Their traffic units are the flow, which limits their abilities towards the variation of the flow size distribution [7]; all of the schemes aim to the bandwidth utilization, which has not directly related to the delay aspect.

The Packet-based schemes [8] [9], on the other hand, choose the path in term of packets. Therefore, packets in the same flow might be transferred via different paths. Least-Loaded First (LLF) [8] [9] is one conventional packet-based scheme. Upon arrival of each packet, the path with the shortest queue would be selected. Compared with Flow-based schemes, the Packet-based schemes use packet as the traffic unit. Since each flow is constituted by one packet at least, the number of packet would be significantly larger than the number of flow, and the available traffic unit of the Packet-based scheme is also far more than that of the Flow-based scheme. Therefore, the Packet-based schemes have sufficient traffic unit to schedule, and then such schemes could efficiently utilize the bandwidth of each path. Nevertheless, most of Packet-based schemes could not guarantee the packet order, and often cause a large packet reordering time [10]. Such disadvantage imposes heavy burden to the router, and are unsuitable for the packet order-sensitive applications such as VoIP or Multimedia streaming.

The Hybrid schemes [11] [12] are the compromised strategies that divide flows into smaller sub-flows in order to increase the available traffic splitting units while preserve the packet order. In FLARE [11], the transmission path for a certain sub-flow is determined by its load and the time interval with the previous one. If the interval is larger than the predefined threshold, such sub-flow could switch the path. LBPF [12] considers the traffic rate of each data flow. Under some specific conditions (i.e., unbalanced condition) the elephant flow (a flow with a high transmission rate and long duration) will change its path so as to mitigate the unevenly distribution of flow size over the paths. The major deficits of the schemes mentioned above are the inabilities of covering both of the packet order and the delay decrease.

Enhanced-Delay Controlled Load Distribution (E-DCLD) [2] is an interesting Packet-based load distribution method. It uses a combination of stochastic delay prediction model and instant queuing status model to simulate the network traffic condition. The major purpose of E-DCLD is to minimize the transmission delay and limit the packet reordering. This scheme is especially applicable for the real-time usages. However, the iterative method, which is used in the scheme, slowly adjusts its traffic splitting ratio in pace with receiving the packets, and such procedure drags on the performance in early stage. Moreover, the more paths available for the system, the more rounds it needs to convergence.

## III. Delay-aimed Load Balancing over Multipath Network

### A. Network structure

We consider a single ingress gateway $G_i$ connected to an egress gateway $G_e$ via $P$ parallel logical paths as shown in Fig. 1. These logical paths could be wired paths, wireless paths, or hybrid paths. When data flow into $G_i$, the load balancing mechanism takes effect. The traffic splitter firstly decides what kind of granularity (i.e., flow, or packet) it would use to control the traffic, then the path selector will choose the path for the current allocating unit according to the policy of the load adapter. We denote the index of each logical path by $p_i \in \{1,2,3,...,P\}$. These logical paths have their respective bandwidths, and $\mu_{p_i}$ represents the bandwidth of the $p_i$-th one, which could be determined by the minimal bandwidth along $p_i$. Besides, each logical path has a fixed delay, i.e. propagation delay, $D_{p_i}$. The fixed delay $D_{p_i}$ mainly depends on the physical conditions and the total length of path $p_i$.

Same as [2], we assume that the input traffic is a combination of Poisson traffic and other traffic with unknown properties, the average arrival rate is $\lambda$. In this model, the load adaptor set in $G_i$ as in Fig. 1 determines the transmission path, on the basis of the traffic arrival rate $\lambda$, bandwidth of each logical path $\{\mu_1,\mu_2,\mu_3,...,\mu_P\}$, and the instantaneous queue size of each path $\{q_1,q_2,q_3,...,q_P\}$.

### B. Total packet delay

The total packet delay is the delay that a packet may experience during the transmission. It mainly consists two parts: The end-to-end delay and the packet reordering delay. The end-to-end delay $d_{p_i}$ is the locally available information and is the sum of the fixed delay and queuing delay:

$$d_{p_i} = D_{p_i} + Q_{p_i}, \tag{1}$$

where $D_{p_i}$ and $Q_{p_i}$ denote the fixed delay and the queuing delay respectively. The queuing delay is the time that a packet will stay in the queue of path $p_i$. The packet reordering delay is mainly determined by the delay gap among the logical paths to a great extent [13]. Therefore, the key problem for the real-time transmission is to low down the end-to-end delay, and at the same time, maintain a small delay gap among these paths.

### C. Problem formulation

The objective of this problem model is to figure out the optimal traffic split ratio, then combine the result with the Surplus Round Robin (SRR) in [14], and eventually minimize total packet delay.

Theoretically, if the input traffic follows Poisson distribution, path $p_i$ is randomly selected with a probability $\psi_{p_i}$ ($0 \leq \psi_{p_i} \leq 1$), and the expected service time for each packet is $1/\mu_{p_i}$, the ingress gateway $G_i$ and its associating logical paths could be regarded as $P \times$ M/M/1 queuing systems. In order to represent both of the theoretical analysis and instantaneous queuing size, we inherit the original combination model to denote the expectation time of a packet stays in the $p_i$ th path $Q_{p_i}$:

$$Q_{p_i} = (1-\omega)\frac{1}{\mu_{p_i} - \psi_{p_i}\lambda} + \omega\frac{q_{p_i}}{\mu_{p_i}}. \tag{2}$$

The first term derives from queuing theory and the second term is the division of the current queue length $q_{p_i}$ over the bandwidth $\mu_{p_i}$. $\omega$ is a weight factor that controls the weight between the theoretical queuing delay and instantaneous queuing delay.

According to Eq. (1) and Eq. (2), the expected end-to-end delay for a packet transmission via path $p_i$ is a function about $\psi_{p_i}$:

$$d_{p_i} = C_{p_i}(\psi_{p_i}) = D_{p_i} + (1-\omega)\frac{1}{\mu_{p_i} - \psi_{p_i}\lambda} + \omega\frac{q_{p_i}}{\mu_{p_i}}. \tag{3}$$

Therefore, the optimization problem can be formulated as follows:

$$Minimize \quad \max_{p_i \in P} C_{p_i}\left(\psi_{p_i}\right), \tag{4}$$

$$subject\ to \quad \sum_{p_i \in P} \psi_{p_i} = 1 \tag{5}$$

$$and \quad 0 \leq \psi_{p_i} \leq \frac{\mu_{p_i}}{\lambda} \leq 1. \tag{6}$$

The second inequality constraint indicates that the rate of arrival traffic should not be larger than the bandwidth of the logical path, otherwise the queue might be infinite long and there is no stable condition of this system.

### D. Problem of the previous E-DCLD algorithm

Although E-DCLD [2] gives a solution of problem 4 (We call the Eq. (4) and its associating constraints problem 4), its initial result is always inaccurate, and the system has to wait several rounds to gradually approach the optimal traffic splitting ratio. The number of rounds depends on the number of logical paths. In reality, such inaccuracies may mislead the path selector in early time. Moreover, if the path status or other parameters are not stable, the path would be under sub-optimal condition in most of the time, which might seriously degrade the system performance. From the result of [15], when the number of path is 5, using Gradually Approaching algorithm (GA), which is the original solution for E-DCLD model, would spend more than 12 rounds to get the optimal distribution ratio. In order to handle the low-convergence rate problem, we propose alternative algorithm in the following section.

## IV. Convex optimization-Based Method (CBM)

### A. Proof of convexity

In this subsection, we prove that the problem 4 is a convex optimization problem. We first prove the convexity of the objective function, then we demonstrate the inequality constraint is convex and equality constraint is affine. The main procedure of the proof is as follows:

**Lemma 1.** *The optimization function is a convex function.*

*Proof:* Firstly, the Hessian matrix of the end-to-end delay function $f = C_{p_i}\left(\psi_{p_i}\right)$ is

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial \psi_1^2} & \frac{\partial^2 f}{\partial \psi_1 \psi_2} & \cdots & \frac{\partial^2 f}{\partial \psi_1 \psi_P} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial \psi_P \psi_1} & \frac{\partial^2 f}{\partial \psi_P \psi_2} & \cdots & \frac{\partial^2 f}{\partial \psi_P^2} \end{pmatrix} \tag{7}$$

Without loss of generality, we consider the first path ($p_i$ = 1), hence

$$H\left(C_1\left(\psi_1\right)\right) = \begin{pmatrix} \frac{2\lambda^2(1-\omega)}{(\mu_1 - \psi_1\lambda)^3} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{8}$$

Because there is only one element in the $P \times P$ matrix, any principle minor of this matrix is non-negative and the matrix $H$ is a positive semi-definite matrix. That means, for each $x$ in the domain of $f$,

$$\nabla^2 f(x) \geq 0. \tag{9}$$

Therefore, the second derivative result of the delay function of each path $C_p\left(\psi_p\right)$ is positive and is a convex function.

On the other hand, the optimization problem 4 could be transformed into the following formation:

$$\left\| C_1\left(\psi_1\right), C_2\left(\psi_2\right), ..., C_P\left(\psi_P\right) \right\|_\infty. \tag{10}$$

The convexity of the optimization function lies on the monotonicity and convexity of the infinite norm, and the convexity of function $C_{p_i}(\psi_{p_i})$. Since the infinite norm is a non-decreasing convex function, and we have already proved the convexity of $C_{p_i}(\psi_{p_i})$, then the composition of the infinite norm and $C_{p_i}(\psi_{p_i})$ is convex. Therefore, the objective function is a convex function. ■

**Lemma 2.** *The equality constraints and inequality constraints are affine and convex functions respectively.*

*Proof:* The equality constraint is 1-norm which is a affine function since for any $x_1, x_2$ in the domain of $\psi_p$, $\|\theta x_1 + (1-\theta)x_2\|_1 = \theta\|x_1\|_1 + (1-\theta)\|x_2\|_1$, and the inequality constraints are the convex function, the constraints of 4 fulfil the requirements. ■

**Theorem 1.** *The objective problem 4 is a convex-optimization problem.*

*Proof:* Lemma 1 proves that the objective function is a convex function; the associating definition domain is a linear transformation on a convex set, and is therefore a convex set; lemma 2 demonstrates the equality constraints and inequality constraints are affine and convex function respectively; the objective problem 4 is a convex optimization problem. ■

## B. Solving procedure

In the remainder of this section, we introduce the proposed Convex-based Method (CBM) to decide the load ratio for each path. Algorithm 1 depicts a framework of the entire procedure. Upon receiving a packet, the method firstly calculate the current rate lambda of the input packets. Besides, the load adapter will get the instant queuing size $q_{p_i}$ from the buffer of each path $p_i$. Such procedure is quite same to the E-DCLD. Since theorem 1 proves that problem 4 is a convex optimization problem, we can use the existed method such as interior point method or Barrier method to get the load distribution vector. Line 4 to line 11 denotes the counting step using interior point method mentioned in [16]. After that, the load adapter will inform the path selector the result. According to the result, the SRR that implemented in path selector will choose the path for the packet.

**Algorithm 1** Convex optimization-Based Method (CBM)

1: **loop**
2: Calculate the data arriving rate $\lambda$ from input traffic.
3: Calculate the instant queuing size $q_{p_i}$ from the buffer for all the paths.
4: Initialize the necessary parameters, including the initial load distribution ratio $\psi_{p_i} = \psi^0_{p_i}$ that satisfies the constraints of such ratio (i.e., $\psi^0_{p_i} = \frac{\mu_{p_i}}{\sum_{p_i \in P} \psi_{p_i}}$ ), the precision control factors, and other parameters used for searching the result.
5: **repeat**
6: Find and update the surrogate duality gap $t$.
7: Based on $t$, count out primal-dual search direction $\Delta D$, according to the pre-defined parameters and the dual functions of objective functions.
8: Use backtrack strategy to find out the step size $s$.
9: Update $\psi_{p_i}$ and other related parameters according to $\Delta D$ and $\psi_{p_i}$.
10: Figure out the related precision.
11: **until** The precision of the current solution satisfied the precision requirement.
12: Set the load distribution ratio $\psi_{p_i}$ for each path.
13: Forward the packet.
14: **end loop**

## V. Performance Evaluation

In this section, we analyse the performances of the proposed Convex-optimization Based Method (CBM) and the Gradually Approaching algorithm (GA). The performance analysis consists of two parts. In the first part, we consider the scenario with the change of offered load. We compare the performances of CBM and GA in terms of the end-to-end packet delay, coefficient variation of end-to-end delay and total packet delay. In the second part, we check the relationship between the number of logical paths and the total delay of the two algorithms.

We suppose the following simulations are conducted under the environment as shown in Fig. 1, which contains a single server, multiple paths and several data flows. The data flow is randomly generated, and coarsely follows the Poisson distribution. The parameter $\lambda$ is proportional to the total bandwidth of the paths, the mean service time of the packets is inversely proportional to the bandwidth capacity. In order to verify whether the algorithms are robust toward the time-varying path status or not, the bandwidth of each logical path is set to be time-varying. Here we use CVX [17] [18], a package for specifying and solving convex programs to solve the Convex-optimization problem.

TABLE I. Major Parameters for Load Ratio-Related Evaluation

| Time($T$) | Path number($P$) | Flow number |
|---|---|---|
| 1 hour | 5 | 5 |
| Input rate | Ratio of load ($\lambda/\mu$) | Omega |
| $R$ × sum(Path rate) | 0.1 ∼0.9 | 0.5 |

## A. Evaluation of various delays versus load ratio

In this simulation, the path rate and fixed delay for each path are random numbers, which are uniformly distributed from 2 to 4. Other major parameters of this simulation are listed in table I.

*1) End-to-end packet delay:* The end-to-end packet delay is the sum of the fixed delay and queuing delay as defined in Eq. (1). Fig. 2 depicts the end-to-end delay of GA and CBM. This figure reveals the following facts: i) As the ratio of input rate to output rate ($\lambda/\mu$) increasing, the mean value of total packet delay rises as well. Because the queuing time $Q_{p_i}$ is likely to increase while the fixed delay $D_{p_i}$ stays the same. ii) When traffic load becomes heavier, the performance gap between GA and CBM gets wider. Because when the traffic load is light, the queue size of each path is relative small and the affect of the changes mainly lies on recounting the path split ratio. When the traffic load is higher, the expectation of queue size is also larger, which makes the initial suboptimal split result of GA become costly. iii) Compared with GA, CBM

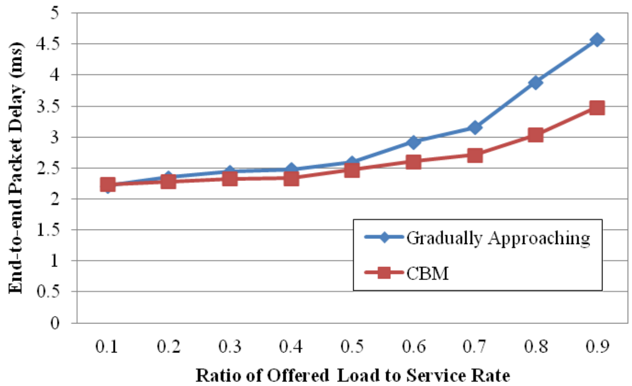


Fig. 2. End-to-end packet delay.

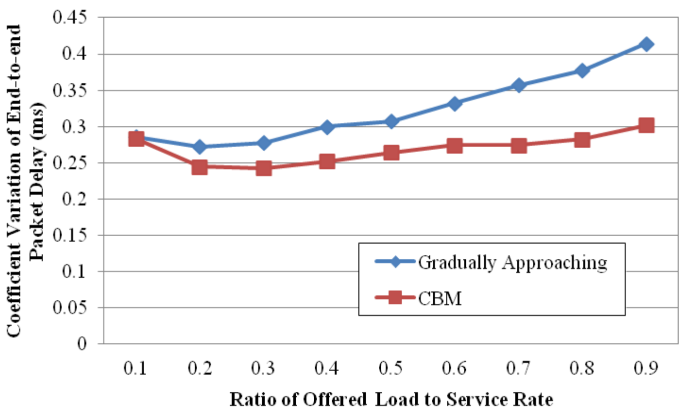


Fig. 3. Coefficient variation of end-to-end packet delay.

has a smaller end-to-end delay and is more suitable for the application that does not require packet order strictly.

*2) Coefficient Variation of end-to-end packet delay:* Fig. 3 depicts the relationship between coefficient variation (CV) of end-to-end packet delay and the ratio of offered load to service rate. A large CV indicates a high probability of packet reordering since the two successive packets may have very different end-to-end delays, and the later packet might be received before the former one. As we can see from Fig. 3, the CV gap between the two algorithms increases along with $\lambda$. The major reason behind such phenomenon is the suboptimal result of GA, and the time-varying paths make the system under the suboptimal condition.

*3) Total packet delay:* The total packet delay is the sum of the end-to-end packet delay and the packet reordering delay. The packet reordering delay might be determined by the upper layer protocol. For simplicity, we suppose the packet reordering delay is the waiting time of a certain packet for the previous packets. The total packet delay is an important indicator for the order-required application. Fig. 4 indicates that when the ratio $(\lambda/\mu)$ is larger than 0.5, the total packet delay counted by GA is pronouncedly larger than CBM. Because both of end-to-end delay and packet reordering delay of CBM are less than GA, the total delay difference between GA and CBM is even more considerable.

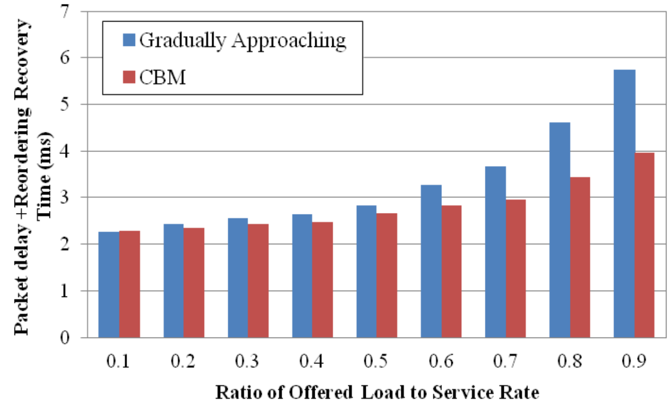


Fig. 4. Total packet delay with time variant path status.

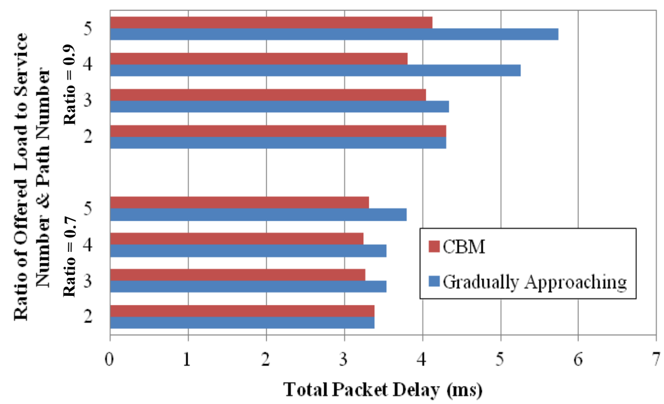


Fig. 5. Total packet delay with different path number and input traffic ratio.

TABLE II. Major Parameters for Path Number-Related Evaluation

| Time($T$) | Path number($P$) | Flow number |
|---|---|---|
| 1 hour | $1 \sim 5$ | $1 \sim 5$ |
| Input rate | Ratio of load $(\lambda/\mu)$ | Omega |
| $R \times$ sum(Path rate) | 0.7, 0.9 | 0.5 |

## B. *Evaluation of total delays versus path number*

In addition to the load ratio-related evaluation, we conduct another simulation, in which the path number is regarded as a variable so as to check the relationship between the delay and the number of paths for GA and CBM. Same as the previous subsection, the initial bandwidth and fixed delay is randomly chosen from 2 to 4. We choose 0.7 and 0.9 as the ratios in the simulation for the sake of distinction. The key parameters are shown in table II.

Although adding a new path means a higher input traffic rate when the load ratio is fixed, the extra logical path enables an extra choice for transmission. Therefore there should not be a dramatic delay variation among different path numbers. However, as we can see from Fig. 5, the larger the number of paths, the bigger delay gap between GA and CBM.

This result testifies our assumption that if the number of paths is large, then GA algorithm might use more rounds to reach the optimal load distribution ratio since it treats only the paths with the largest delay and smallest delay. Meanwhile, the CBM could maintain a consistent performance regardless the number of paths.

### C. Summary

The experiments above demonstrate the advantages of the proposed CBM algorithm. CBM can avoid the weakness of low-convergence rate of GA algorithm for E-DCLD model when path status is changing with time, and achieve a low-delay split strategy. Besides, different from GA, the number of paths does not have much impact on the result of CBM.

## VI. Conclusion

Multipath routing is regarded as a promising technique due to its effective usage toward the available multiple paths. To this end, a large body of literatures has been devoted in this area. Many techniques developed in previous researches focus on the bandwidth utilization and packet order reservation, and are not well suited for real-time applications. The E-DCLD raises a delay-related mechanism. It can meet the requirements of delay and packet order. However, the solution for E-DCLD is undesirable for unstable path scenario, because its initial result is not desirable and need time to convergence to the optimal one. To address the inefficiencies of the original solution, we prove the convexity of the model and propose a Convex optimization-Based Method (CBM), which could count out the optimal traffic split ratio without redundant convergence round. Experimental results demonstrate that our scheme can significantly decrease the end-to-end packet delay and total packet delay.

To further expand upon the analysis and load balancing problems, the next step is to study the multi-source and multi-destination scenario, and propose a solution to meet the fairness requirement among different data flows as well as minimize the transmission delay.

## VII. Acknowledgement

This work was conducted under the national project, "Research and development of Movable ICT-Units for emergency transportation into disaster-affected areas and multi-unit connection", supported by the Ministry of Internal Affairs and Communications (MIC), Japan.

## References

[1] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture," RFC 3031 (Proposed Standard), Internet Engineering Task Force, Jan. 2001, updated by RFCs 6178, 6790. [Online]. Available: http://www.ietf.org/rfc/rfc3031.txt

[2] S. Prabhavat, H. Nishiyama, N. Ansari, and N. Kato, "Effective delay-controlled load distribution over multipath networks," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 22, no. 10, pp. 1730–1741, 2011.

[3] A. Zinin., "Cisco ip routing, packet forwarding and intra-domain routing protocols." *MA: Addison Wesley*, 2002.

[4] D. Thaler and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection," RFC 2991 (Informational), Internet Engineering Task Force, Nov. 2000. [Online]. Available: http://www.ietf.org/rfc/rfc2991.txt

[5] C. Villamizar, "OSPF Optimized multipath(OSPF-OMP)," IETF Internet draft, work in progress, Internet Engineering Task Force, Feb. 1999.

[6] C. Hopps, "Analysis of an Equal-Cost Multi-Path Algorithm," RFC 2992 (Informational), Internet Engineering Task Force, Nov. 2000. [Online]. Available: http://www.ietf.org/rfc/rfc2992.txt

[7] Z. Cao, Z. Wang, and E. Zegura, "Performance of hashing-based schemes for internet load balancing," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1, 2000, pp. 332–341.

[8] C.-C. Hui and S. Chanson, "Hydrodynamic load balancing," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 10, no. 11, pp. 1118–1137, 1999.

[9] K. Shin and C.-J. Hou, "Design and evaluation of effective load sharing in distributed real-time systems," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 5, no. 7, pp. 704–719, 1994.

[10] N. M. Piratla and A. Jayasumana, "Reordering of packets due to multipath forwarding - an analysis," in *Communications, 2006. ICC '06. IEEE International Conference on*, vol. 2, 2006, pp. 829–834.

[11] S. Kandula, D. Katabi, S. Sinha, and A. Berger, "Dynamic load balancing without packet reordering," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 2, pp. 51–62, Mar. 2007. [Online]. Available: http://doi.acm.org/10.1145/1232919.1232925

[12] W. Shi, M. MacGregor, and P. Gburzynski, "Load balancing for parallel forwarding," *Networking, IEEE/ACM Transactions on*, vol. 13, no. 4, pp. 790 – 801, Aug. 2005.

[13] S. Prabhavat, H. Nishiyama, N. Ansari, and N. Kato, "On the performance analysis of traffic splitting on load imbalancing and packet reordering of bursty traffic," in *Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on*, 2009, pp. 236–240.

[14] H. Adiseshu, G. Parulkar, and G. Varghese, "A reliable and scalable striping protocol," in *Conference proceedings on Applications, technologies, architectures, and protocols for computer communications*, ser. SIGCOMM '96. New York, NY, USA: ACM, 1996, pp. 131–141. [Online]. Available: http://doi.acm.org/10.1145/248156.248169

[15] S. Prabhavat, H. Nishiyama, Y. Nemoto, and N. Kato, "Load distribution with queuing delay bound over multipath networks: Rate control using stochastic delay prediction," *26th International Communications Satellite Systems Conference (ICSSC)*, 2008.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[17] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx, Sep. 2012.

[18] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/ boyd/graph_dcp.html.