REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
08/17/2020	Final Technical Report	06/01/2017 - 05/30/2020

4. TITLE AND SUBTITLE
SDN-Enabled Time-Sensitive Network Reconfiguration

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-17-1-2419

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kevin Tang

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cornell University
OFFICE OF SPONSORED PROGRAMS
373 PINE TREE RD
ITHACA NY 14850-2820

8. PERFORMING ORGANIZATION REPORT NUMBER
Cornell OSP 81783

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
MATH COMPUTERS & INFO RESEARCH DIV
875 N. Randolph Street
Arlington VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; Distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
A general model and its solution are developed to determine how network routing can be reconfigured quickly without incurring transient congestion. Assuming both initial and target configurations are congestion-free, it is known that transient congestion may still occur during the reconfiguration process if links contain a mix of traffic flows following old and new routing rules, resulting from variation of switch reaction time and propagation delay differences among paths. We consider these factors by explicitly incorporating timing uncertainty intervals into the model. The model leads to an optimization problem whose solution represents a fast (in terms of actual physical time) congestion-free routing reconfiguration.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Ao Kevin Tang / **19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU		607-255-4803

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATE COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33315-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report. e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Final Technical Report for "SDN-Enabled Time-Sensitive Network Reconfiguration" (N00014-17-1-2419)

PI: Ao Kevin Tang

Institution: Cornell University

Abstract:

A general model and its solution are developed to determine how network routing can be reconfigured quickly without incurring transient congestion. Assuming both initial and target configurations are congestion-free, it is known that transient congestion may still occur during the reconfiguration process if links contain a mix of traffic flows following old and new routing rules, resulting from variation of switch reaction time and propagation delay differences among paths. We consider these factors by explicitly incorporating timing uncertainty intervals into the model. The model leads to an optimization problem whose solution represents a fast (in terms of actual physical time) congestion-free routing reconfiguration. Our formulation naturally reduces to existing work of finding minimal number of algorithmic update steps when the timing uncertainty intervals are very large, meaning we have little prior knowledge about them. The optimization problem is shown to be a Mixed Integer Linear Program (MILP) with a polynomial-size constraint set, and is proved to be NP-hard. We then further introduce an approximation algorithm with performance guarantee to solve the problem efficiently. In particular, it is demonstrated that timing information can possibly accelerate the update process, even if more steps are involved.

More specifically, we proceed to do the following three tasks. We characterize the optimal solution itself. This includes both optimality condition and complexity of the solution. We are also investigated tighter approximation algorithms as well as other fast effective heuristics. We then develop a model of network traffic dynamics. The model includes two independent variables: user input rate (demand) and router split ratio. There are two main features of this model: careful description of the varying queueing delay and spatial heterogeneity of a flow's rate due to different propagation delay. Both are necessary details (though typically ignored in existing literature) for predicting fast network dynamics. Finally, we investigate and compare various network architecture choices such as, for example, distributed vs centralized routing control.

1. Description of Topics and Results

Routing Stability in Hybrid Software-Defined Networks

Software-defined networks (SDNs) facilitate more efficient routing of traffic flows using centralized network view. On the other hand, traditional distributed routing still enjoys the advantage of better scalability, robustness, and swift reaction to events such as failure. There are therefore significant potential benefits to adopt a hybrid operation where both distributed and centralized routing mechanisms co-exist. This hybrid operation however imposes a new challenge to network stability since a poor and inconsistent design can lead to repeated route switching when the two control mechanisms take turns to adjust the routes. We have discussed ways of solving the stability problem. We first define stability for hybrid SDNs and then establish a per-priority stabilizing framework to obtain stable routing patterns. For each priority class, we discuss three approaches to reach hybrid SDN stability: global optimization, greedy, and local search. We find that the proposed local search provides the best tradeoff among cost performance, computational complexity, and route disturbance. Furthermore, we design a system on a centralized controller, which utilizes those algorithms to stabilize the network. The design is implemented and extensively tested by simulations using realistic network information, including a trace of the Abilene network and data from a tier-1 Internet service provider's backbone network.

Analysis of Fast Adaptive Traffic Engineering Using a Feedback Control Model

We analyze fast adaptive traffic engineering from a feedback control perspective. We provide a model that characterizes the system parameters' effects on the performance of the dynamic routing system. This allows quantitative analysis of adaptive traffic engineering control laws and its design parameter choices. We then specialize the general framework in two representative network topologies and derive the stability conditions for their dynamic routing systems. Experiments are carried out to compare against our theoretical predictions considering two forms of adaptive traffic engineering control law as examples. Together they provide systematic insights on the relations among several network factors and the intrinsic tradeoff among different network control objectives.

Accurate Rate-Aware Flow-level Traffic Splitting

We develop a technology to accurately realize given traffic split ratios in switches with small performance degradation. For given traffic split ratios calculated mathematically by TE algorithms in the control plane, the load distribution mechanisms in the data plane implement such splits without breaking flows. Treating all flows equally, the state-of-the-art approaches deployed in switches do not provide enough accuracy especially when facing non-uniform flow size distribution. We instead propose a dynamic load distribution scheme based on the collected load sharing statistics. It finds the most accurate traffic splits with minimum route changes. We implement our solution in Open vSwitch (OVS).

Trace-driven and end-toend experiments demonstrate that 1) our approach effectively adjusts load distribution in real time to mitigate the inaccuracy of splits caused by the variation of flow size distribution, 2) it outperforms the existing approaches with respect to both higher accuracy and lower level of route changes, and 3) it requires path changes for less flows when routing strategies are reconfigured, hence leads to better flow experience such as higher goodput.

Fast In-Network Congestion Management

WANs are often over-provisioned to accommodate worst-case operating conditions, with many links typically running at only around 30% capacity. In this project, we show that in-network congestion management can play an important role in increasing network utilization. To mitigate the effects of in-network congestion caused by rapid variations in traffic demand, we propose using high-frequency traffic control (HFTraC) algorithms that exchange real-time flow rate and buffer occupancy information between routers to dynamically coordinate their link-service rates. We show that the design of such dynamic link-service rate policies can be cast as a distributed optimal control problem that allows us to systematically explore an enlarged design space of in-network congestion management algorithms. This also provides a means of quantitatively comparing different controller architectures: we show, perhaps surprisingly, that centralized control is not always better. We implement and evaluate HFTraC in the face of rapidly varying UDP and TCP flows and in combination with AQM algorithms. Using a custom experimental testbed, a Mininet emulator, and a production WAN, we show that HFTraC leads to up to 66% decreases in packet loss rates at high link utilizations as compared to FIFO policies.

3. Student Training

Several results from this project have been included in a graduate course (ECE 5800: Control and Optimization of Information Networks) at Cornell. The grant also partially supported four PhD students (Shih-Hao Tseng, Yingjie Bi, Ning Wu and Jiangnan Cheng). Shih-Hao Tseng graduated in 2018 and is a now a postdoc at Caltech. Ning Wu graduated in 2018 and is currently with MODE group. Yingjie Bi graduated in 2020 and is now a postdoc at UC Berkeley. Jiangnan Cheng will graduate in 2022.

4. Results Dissemination

In terms of publications, we have published most related results, including seven Conference papers and three journal papers. In terms of talks, besides several conference presentations, we have also given several invited seminars at both universities and industry labs.

5. Awards and Honors

Yingjie Bi: Best PhD thesis award, ECE, Cornell, 2020

Kevin Tang: Senior Member, Association for Computing Machinery (ACM), 2019

Kevin Tang: IBM Faculty Award, IBM, 2018

The whole team: Winner of the AT&T SDN Network Design Challenge, 2017

6. Future Direction

One of my future research directions is to automatic network management in computer networks. It will be a critical technology to truly realize a flexible high performance network. I plan to shift my focus more from the core to the edge. I hope to be able to contribute to the emergence of Telco cloud, in particular to enable it to quickly assemble the right amount of network resources for different applications. I believe one key ingredient of that technology is the ability to reconfigure network flow paths in a timely fashion. The results from the current project should serve as an excellent basis for my new endeavor.

7. Main Related Publication

[1] C. Lim, K. Lee, H. Wang, H. Weatherspoon and A. Tang, "Packet Clustering Introduced by Routers: Modeling, Analysis and Experiments", *ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 4(3), 15:1-28, September,* 2019

[2] S. Tseng, A. Tang, G. Choudhury and S. Tse, "Routing Stability in Hybrid Software-Defined Networks", *IEEE/ACM Transactions on Networking*, 27(2): 790-804, April 2019.

[3] N. Wu, Y. Bi, N. Michael, A. Tang, J. Doyle and N. Matni, "A Control-Theoretic Approach to In-Network Congestion Management", *IEEE/ACM Transactions on Networking*, 26(6): 2443-2456, December 2018.

[4] J. Huang, Q. Cai, R. Agarwal and A. Tang, "I10: A Remote Storage I/O Stack for High-Speed Links", Proceedings of 17th USENIX NSDI 2020.

[5] J. Cheng, J. Huang, and A. Tang, "Acceleration of Multipath TCP by Letting ACKs Take the Shortest Path", Proceedings of 40th IEEE Sarnoff Symposium, 2019.

[6] Y. Bi and A. Tang, "Uncertainty-Aware Optimization for Network Provisioning and Routing", Proceedings of CISS, 2019.

[7] S. Tseng and A. Tang, "Coflow Deadline Scheduling via Network-Aware Optimization", Proceedings of Annual Allerton conference, 2018

[8] N. Wu, S. Tseng and A. Tang, "Accurate Rate-Aware Flow-Level Traffic Splitting", Proceedings of Annual Allerton conference, 2018

[9] N. Wu and A. Tang, "Analysis of Fast Adaptive Traffic Engineering Using a Feedback Control Model", Proceedings of ACC, 2018

[10] N. Wu and A. Tang, "End-to-end Network Throughput Optimization Through Last-mile Diversity", Proceedings of CISS, 2018