On Quantifying Performance Enhancement of Distributed SDN Architecture

Ziyao Zhang
Imperial College London
London, United Kingdom
ziyao.zhang15@imperial.ac.uk

Liang Ma
IBM T. J. Watson Research Center
Yorktown Heights, NY, United States
maliang@us.ibm.com

Kin K. Leung
Imperial College London
London, United Kingdom
kin.leung@imperial.ac.uk

***Abstract*—We study and quantify the performance enhancement achievable under the distributed software-defined networking (SDN) architecture, which is the composition of interconnected SDN-based network domains. Under a generic network model, we develop mathematical frameworks to analyze the performance metric of distributed SDN under various inter-domain synchronization levels and network properties. Simulations based on both synthetic and real network topologies confirm the accuracy of our theoretical results. To the best of our knowledge, this is the first work quantifying the performance of distributed SDN analytically, which provides fundamental guidance for future SDN protocol designs and performance estimation.**

I. Introduction

Distributed software-defined networks (SDN), consisting of multiple inter-connected network domains, each managed by one SDN controller, is an emerging networking architecture that offers balanced centralized control (for intra-domain communications) and distributed operations (for inter-domain communications). Under such networking paradigm, most existing works focus on designing complicated controller-synchronization strategies to improve the joint controller-decision-making for inter-domain routing. However, there is still a lack of fundamental understanding of how the performance of distributed SDN is related to network attributes, thus impossible to justify the necessity of complicated strategies. In this regard, we analyze and quantify the performance enhancement of distributed SDN architectures, influenced by network topological properties and inter-domain synchronization levels.

In this paper, we employ the average path length (APL), measured by the average number of hops of the *shortest* path connecting two arbitrary nodes in a network, as the performance metric of interest. Under a generic network model, we aim to answer two critical questions. (i) What is the achievable performance when each domain operates independently, i.e., no inter-domain synchronizations? (ii) How does the complete inter-domain synchronization, i.e., each domain has the full knowledge of other domains' current status, assist in the performance enhancement? We answer these questions by developing a mathematical framework and prove that APL is a logarithmic function of network settings. On top of these two extreme cases, we further investigate to what extent partial inter-domain synchronizations may help reduce the APL under one basic flow construction scheme. The accuracy of these analytical results are then confirmed by extensive evaluations under both synthetic and real networks.

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

II. Generic Network Model

To understand the performance of the distributed SDN architecture, we formulate the network as an undirected graph according to a two-tier hierarchical model. Specifically, tier-1 consists of m independent domains, each modeled as an undirected random graph with n nodes that are interconnected following a specific degree distribution extracted from real networks (i.e., there are mn nodes in total). Then abstracting each domain as a single vertex, we connect these vertices in tier-2 according to a domain-wise degree (i.e., the number of neighboring domains for a domain) distribution, referred to as the *domain-wise network*, denoted by graph $\mathcal{G}_d = (V, E)$ ($|V| = m$). The existence of an edge in E connecting two vertices in V implies that the two network domains corresponding to these two abstracted vertices are interconnected. Then edges in E are mapped to physical links in the following way: For each $e \in E$ with end-points corresponding to tier-1 domains $\mathcal{A}_i$ and $\mathcal{A}_j$, (i) randomly select two nodes w_1 from $\mathcal{A}_i$ and w_2 from $\mathcal{A}_j$ and connect these two nodes if edge $w_1 w_2$ does not already exist; (ii) repeat such edge construction between $\mathcal{A}_i$ and $\mathcal{A}_2$ β times. Since the same node pairs may be selected in this edge construction process, β therefore represents the largest number of edges between two domains.

III. Performance Quantifications

A. No Inter-domain Synchronization

The worst case of the distributed SDN happens when there is no inter-domain synchronizations, for which the network is operated by BGP-like protocols. Given the degree distribution in each domain, let z_i denote the expected number of nodes that are i-hop away from an arbitrary node. Then the APL l between two arbitrary nodes in a domain with n nodes is $l \simeq ln(n/z_1)/ln(z_2/z_1) + 1$ according to [1]. Similarly, let z'_i be the number of vertices which are i-hop away from an arbitrary vertex in the domain-wise network. Let Y be the random variable of the number of domains that the shortest path between two arbitrary nodes traverses. Then $\Delta := \mathbb{E}[Y] \simeq ln(m/z'_1)/ln(z'_2/z'_1) + 2$. In addition, motivated by the methodology in [1], we derive that under BGP-like protocols, w.r.t. two arbitrary nodes residing in two different domains, the average length of path segments l' in a domain that does not include the destination node is

$$l' \simeq \begin{cases} \frac{n-\gamma}{n}\left(\frac{ln(\frac{n+1-\gamma}{\gamma})}{ln(z_2/z_1)}+1\right) & \text{for } \gamma \leq (n+1)/2, \\ \frac{n-\gamma}{n} & \text{for } \gamma > (n+1)/2, \end{cases} \quad (1)$$

where $\gamma = n(1-(1-1/n)^{\beta})$. Then the APL under BGP for two arbitrary nodes in different domains is $L_{\text{BGP}} \simeq (l'+1)(\Delta-1)+l$.

B. Complete Synchronization Among Domains

To compute the APL under complete synchronization among domains, we consider a specific family of our distributed SDN network model, where the domain-wise network forms a bus topology, i.e., a special tree with vertex degree being at most 2, referred to as *bus network*. In a bus network with k domains, we sequentially label the domains from one end to the other end as $\mathcal{A}_1, \ldots, \mathcal{A}_k$. Then randomly select a node from $\mathcal{A}_1$ and another node from $\mathcal{A}_k$, and let random variable D_k denote the distance of the shortest path between these two different nodes and the corresponding probability mass function (PMF) $f_{D_k}(d)$. It is trivial when $k=1$, $f_{D_1}(d) = \Pr(D_1 = d) = z_d/n, d = 0,1,2,\ldots$. When $k>1$, we can leverage network structures to derive

$$f_{D_k}(d) = \begin{cases} (1-F_U(d-1))^{\beta^{k-1}} & d \geq k, \\ -(1-F_U(d))^{\beta^{k-1}} & \\ & \\ 1-(F_U(d))^{\beta^{k-1}} & d = k-1, \end{cases} \quad (2)$$

where $F_U(d)$ is the Cumulative Distribution Function (CDF) of random variable $U := k-1+\overbrace{D_1+\cdots+D_1}^{k \text{ RVs}}$. Then the expected value of U is $L_k := \mathbb{E}[U]$. Moreover, we prove the following theorem.

Theorem 1. *$L_k < L_{k+1}$ under the 2-tier network model.*

Theorem 1 implies that under the 2-tier network model, on average, the shortest path between two arbitrary nodes also traverses the minimum number of domains, i.e., traversing additional domains does not help in reducing APL. As such, let $h_Y(y)$ denote the PMF of Y (section III-A). Then the APL L^* under complete inter-domain synchronizations for two arbitrary nodes in different domains is $L^* = \sum_{y=2}^{m} L_y h_Y(y)$.

C. Partial Synchronization Among Neighboring Domains

Based on the previous two extreme inter-domain synchronization cases, we now consider how minimum inter-domain synchronizations can improve the performance of distributed SDN. Specifically, suppose each domain only synchronizes with its neighboring domains, i.e., each domain knows and only knows the routing information of all its neighboring domains. With this information, we quantify the performance of a basic SDN-based flow construction strategy, as stated below.

1) The constructed path traverses the same sequence of domains as that selected by BGP-like protocols, and the path is jointly constructed by each controller in these domains;
2) Upon a flow request, the SDN controller in the current domain follows the instruction from the previous domain(s); if no such instruction exists, go to 3);
3) Viewing all synchronized domains as a whole, the SDN controller in the current domain selects a path starting from the ingress node to the closest egress node, which connects to unsynchronized domains.

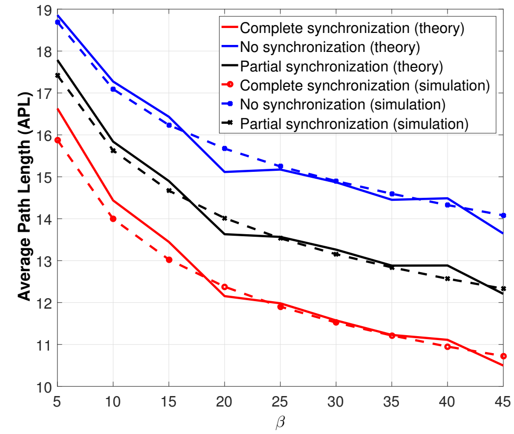

Fig. 1. Analytical versus simulation results under three synchronization levels.

Similar to the definition of L_k in Section III-B, let L_k^{SDN} be the average path length in a bus network with k domains under the above SDN-based flow construction strategy. Then

$$L_k^{\text{SDN}} = \begin{cases} (\frac{k}{2}-1)L_{\text{unit}} + L_2 + \frac{k}{2} - 1 & k \text{ is even}, \\ & \\ \frac{k-1}{2}L_{\text{unit}} + l + \frac{k-1}{2} & k \text{ is odd}, \end{cases} \quad (3)$$

where $k \geq 2$, and L_{unit} is the path length incurred by step 3). Then the APL under the above SDN-based flow construction strategy L_{SDN} is $L_{\text{SDN}} = \sum_{y=2}^{m} L_y^{\text{SDN}} h_Y(y)$.

IV. Evaluations and Conclusions

For evaluations, the intra-domain degree distribution follows the topologies collected from the Rocketfuel Project; while for the domain-wise network, we use artificially generated degree distribution. A graph with a specific degree distribution is realized in the following way: Randomly connect two vertices if they are neither connected already nor exceed the degree requirement; repeat this process until the targeted degree distribution in the graph is achieved.

The evaluation results are reported in Fig. 1, where we compare the simulation and numerical results using our analytical framework under three different inter-domain synchronization levels. Fig. 1 confirms that the numerical results obtained from our analytical framework achieves high accuracy comparing to the simulation results under all cases. More importantly, Fig. 1 reveals that the basic SDN-based strategy is able to reduce the gap to the optimal value (i.e., corresponding to complete inter-domain synchronizations) by around 50%, which implies that even the minimum inter-domain synchronizations can significantly improve the network performance.

References

[1] M. E. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Physical Review E*, vol. 64, no. 2, p. 026118, 2001.