On Quantifying Performance Enhancement of Distributed SDN Architecture

Ziyao (Spike) Zhang
Imperial College London

Liang Ma
IBM US

Kin K. Leung
Imperial College London

September 2017
DAIS-ITA Annual Fall Meting

Main Contents

- Motivation
- Problem statement
- Network Model
- Analysis of 3 scenarios correspond to different synchronization levels
- Evaluation and results
- Summary
- Future plans

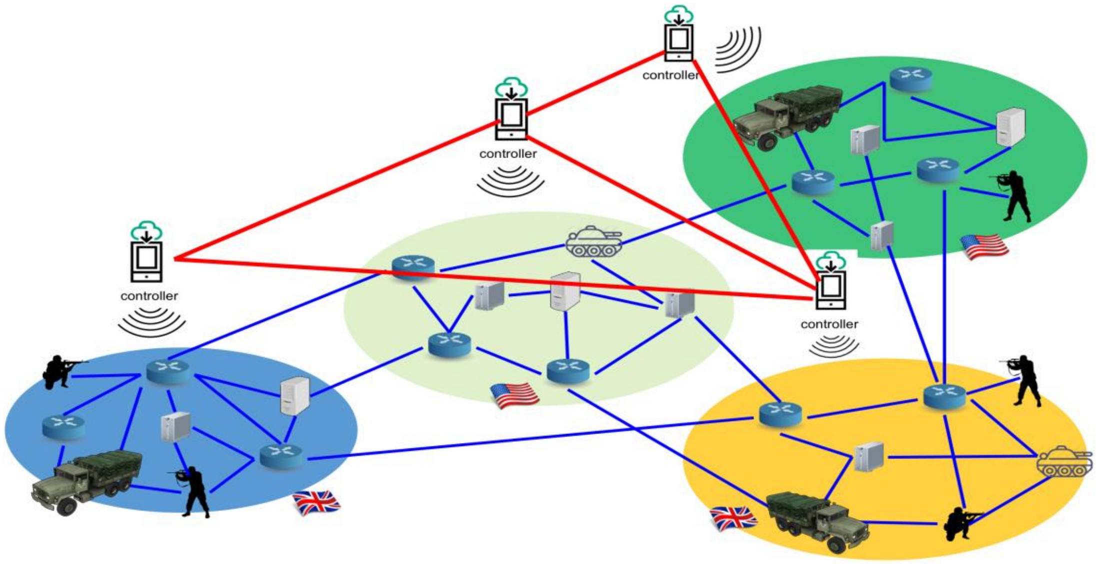
Motivation

The SDN paradigm offers potential performance enhancement of a network

Many efforts from both academia and industry for the design of protocols/architectures to take advantage of SDN

Existing work concerning performance evaluation are mainly prototyping/emulation based

Problem: the lack of fundamental understanding of the bounds of performance enhancement of SDN



Problem statement

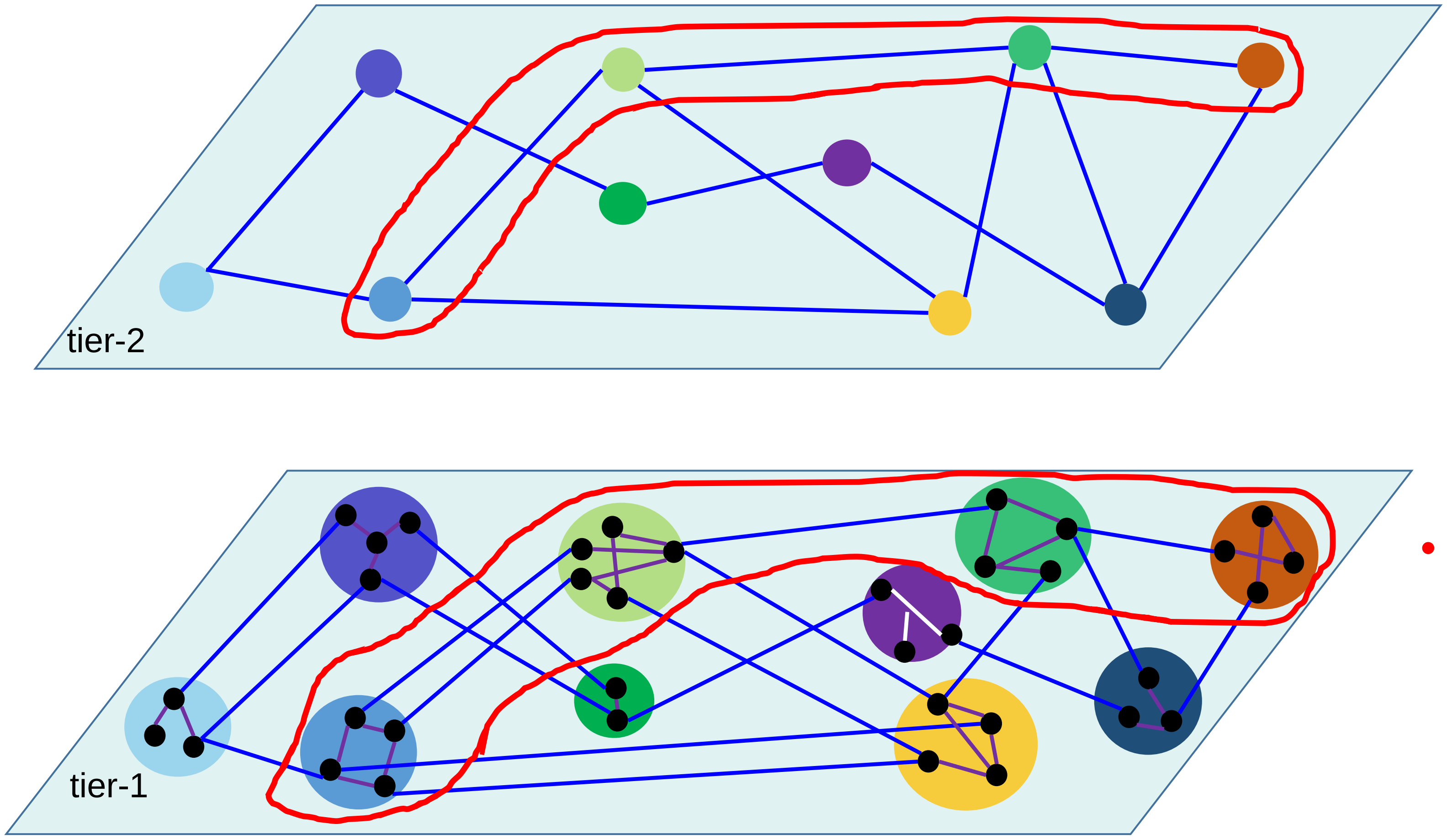
- Goal: gain fundamental (analytical) understandings of performance enhancement offered by SDN
- Performance metric: average path length (APL) measured by # hops, will be extended to include non-uniform edge weight scenarios
- Methodology:

(1) Construct a network using a generic model for analysis

(2) Propose mathematical models to analyse APL under different controller synchronization levels:
  - optimal performance: complete control plane syncs
  - worst performance: no syncs among domains
  - somewhere in the middle: partial syncs via SDN controllers

(3) Simulation confirming the accuracy of proposed analytical model

Network Model



Tier-1: domains whose topology is decided by degree distribution extracted from real network

Tier-2: each domain is abstracted as a single node and two domains are joined by a link in domain-wise topology if there are physical connections

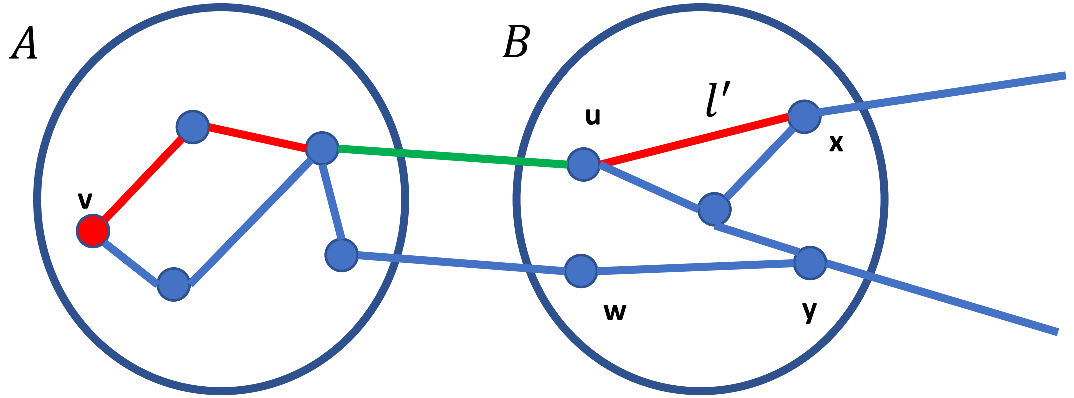
- *Parameters for the network model*
  - n: # nodes in one domain
  - m: # domains in one domain
  - β : max # inter-domain connections between two domains
  - γ: # gateway nodes in one domain for connection with another domain
  - Inter- and inter- domain degree distribution

This model reflects the nature of the distributed SDN network on page 3, and it is more generic compare to our old network model presented in June

Network model (Cont.)

- Controller synchronization: when two controllers synchronize, they share with each other the distance between desired ingress/egress node pairs

Example: if controller A syncs with controller B, then A will know distance between u/x,y and w/x,y, and vice versa



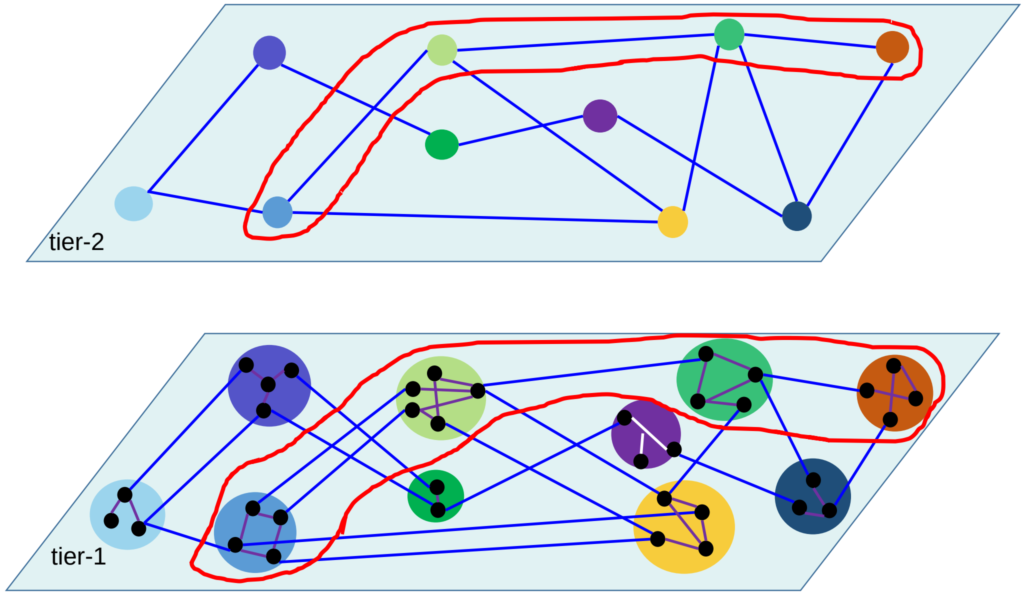
- Controller placement/organization: no specific requirements on the location of controllers or how they are organized. The only assumption is that each domain has one logically centralized controller

Recall: our goal is to obtain fundamental understandings of performance enhancement by SDN controllers, not where or how they are placed

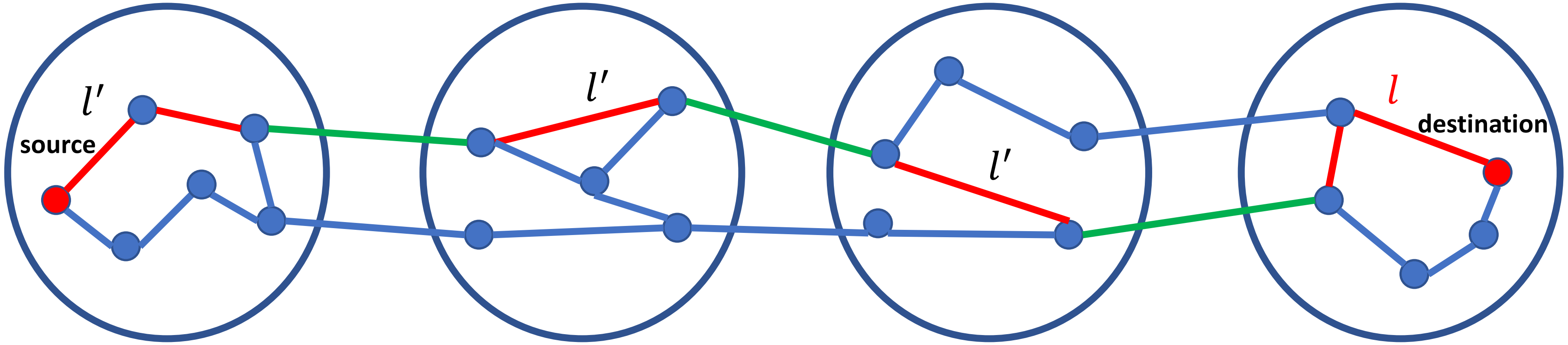
Inter-domain routing strategies - no sync among domains

- 1. based on the addresses of src/dest, decide a domain-wise route by choose the shortest
- 2. find nearest gateway node in current domain from current sending node and egress the packet to the next domain on domain-wise route

Assumptions: tier-2 topology known by all domains (as in BGP)



BGP-like protocol
(Border Gateway Protocol)



Main analytical results- no sync among domains

High level goal: L_{BGP}= *APL in one domain × avg. # domains on a domain-wise route*

- $l \simeq ln(n/z_1)/ln(z_2/z_1) + 1$ **(1)** **l: APL between two random nodes within one domain**
- $\Delta \simeq ln(m/z'_1)/ln(z'_2/z'_1) + 2$ **(2)** **Δ: avg. # domains on a domain-wise route**
- $$l' \simeq \begin{cases} \frac{n-\gamma}{n}\left(\frac{ln(\frac{n+1-\gamma}{\gamma})}{ln(z_2/z_1)} + 1\right) & \text{for } \gamma \le (n+1)/2, \\ \frac{n-\gamma}{n} & \text{for } \gamma > (n+1)/2. \end{cases} \quad (3)$$
- $\gamma = n(1-(1-1/n)^{\beta})$ **(4)** **γ: # gateway nodes in one domain**
- $L_{\text{BGP}} \simeq (l'+1)(\Delta - 1) + l$ **(5)** **L_{BGP} : APL under BGP**

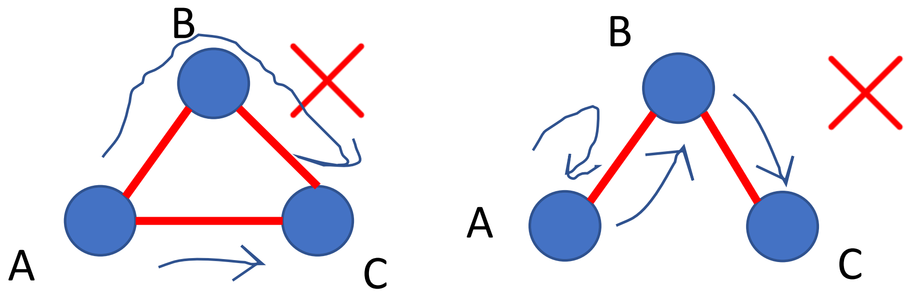
z_i: avg. # vertices i hops away from an arbitrary node
z'_i: corresponding z_i in domain-wise network
l': average distance between an ordinary node and its nearest gateway node

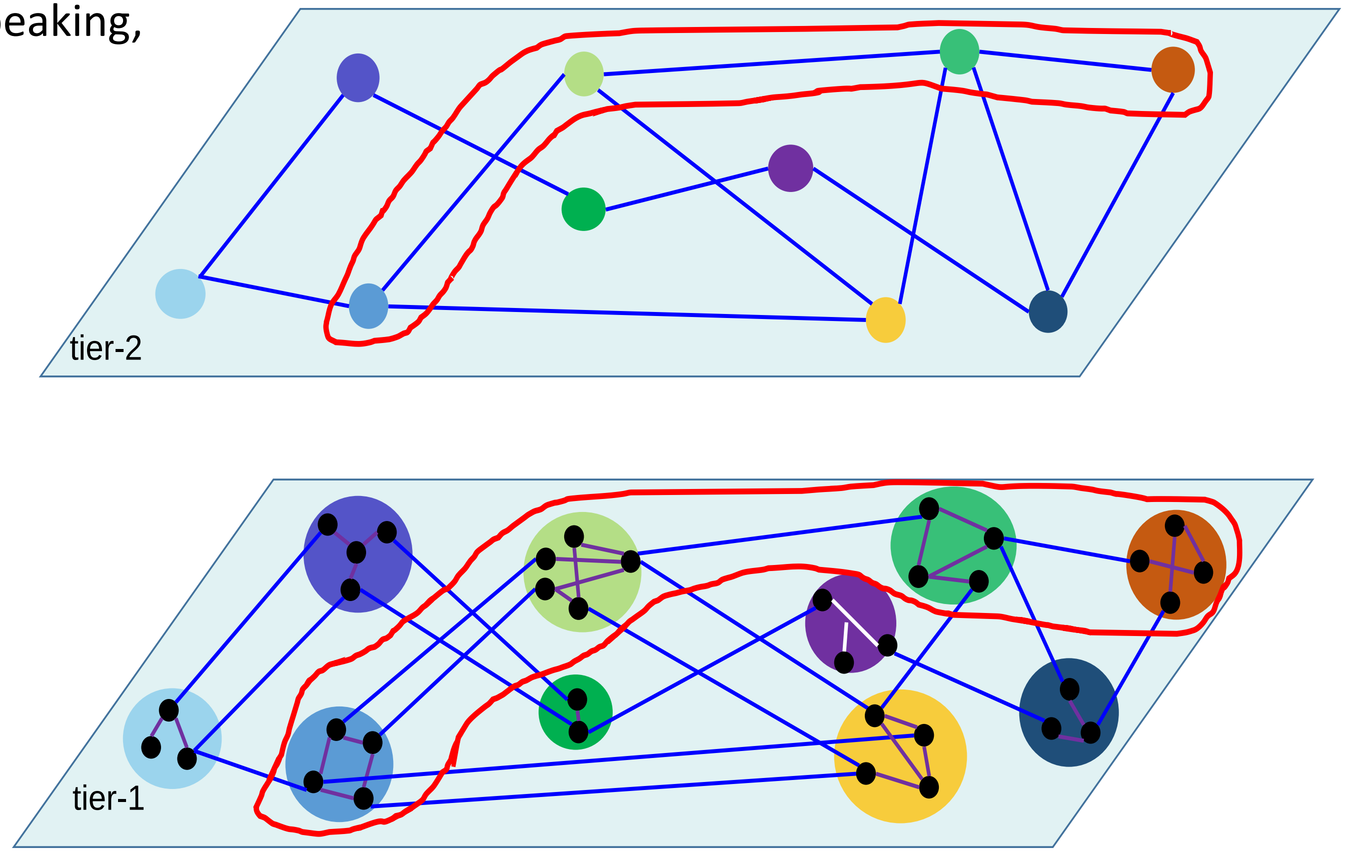
- equation (1): a result drawn from existing literature
- equation (2): equation (1) applied to domain-wise topology
- equation (3): we extend the analysis of shortest APL between two nodes into the analysis of shortest APL between one node and a set of nodes

Inter-domain routing strategies - complete sync among domains

Observations: any domain-wise routes for a src/dest pair will involve domains connected in a bus topology from the source domain to the destination domain

some scenarios (below) may occur. But statistically speaking, these will lead to a larger expectation of APL



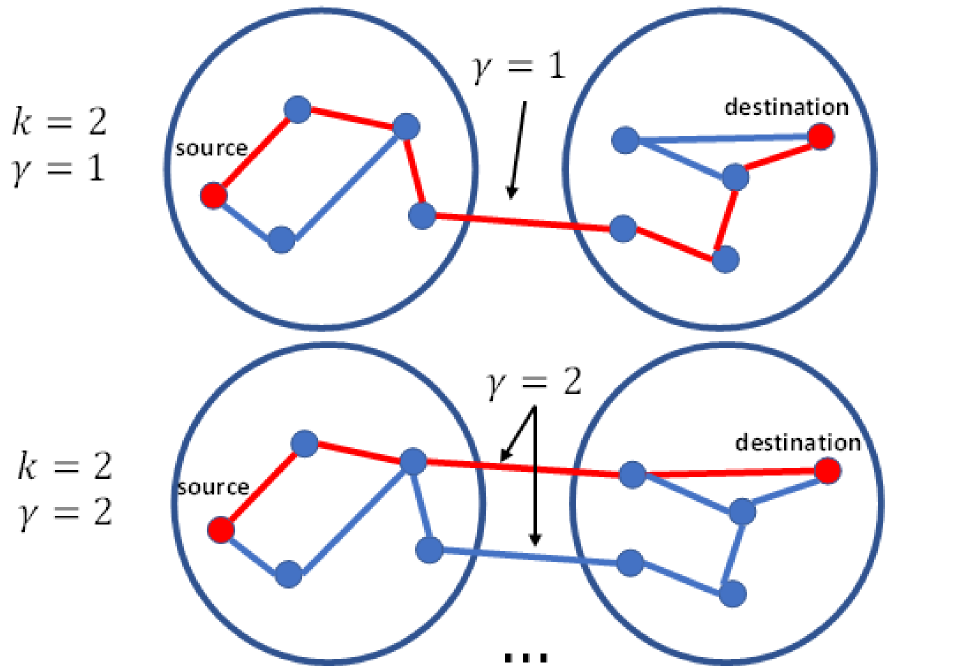


- **Theorem 1: on average, an optimal route involves minimum number of domains. (under our 2-tier network model)**

- **Based on theorem 1, we need to know: 1) APL between two nodes in the first and last domains of a bus; 2) the distribution of the length of the bus**

Main analytical results - complete sync among domains

Calculate the APL between two nodes in the first and last domain of domains connected in a bus topology



APL between two arbitrary nodes in two synced domains

Main elements of the developed analytical framework:

- For individual domains: (step 1)
  - Input: degree distribution
  - Output: distance distribution
- For domains connected in a bus: (step 2)
  - inputs: (1) distance distribution in each domain
    (2) network parameter β
  - Output: expectation of the APL between two arbitrary nodes in the first and last domains in the bus, with respect to a specific β
- For tier-2 topology: (step 3)
  - Input: domain-wise degree distribution
  - Output: distribution of the length of the bus

Main analytical results - complete sync among domains (Cont.)

Inputs: (1) distance distribution in each domain (step 1); (2) network parameter β (step 2);
(3) domain-wise distance distribution (step 3)

Output: APL between two randomly selected nodes in two domains

- $f_{D_1}(d) = \Pr(D_1 = d) = z_d/n, d = 0, 1, 2, \ldots$
- $f_{D_k}(d) = \begin{cases} (1 - F_U(d-1))^{\beta^{k-1}} - (1 - F_U(d))^{\beta^{k-1}} & d \geq k, \\ 1 - (F_U(d))^{\beta^{k-1}} & d = k-1. \end{cases}$
- $L_k := \mathbb{E}[D_k]$
- $L^* = \sum_{y=2}^{m} L_y h_Y(y)$

L_k: mean of RV D_k

L^*: APL under complete inter-domain synchronizations

$f_{D_1}(d)$: distance distribution in one domain

$f_{D_k}(d)$: joint distance distribution between two random nodes in the first and last domains of k domains connected in a bus topology, with network parameter β

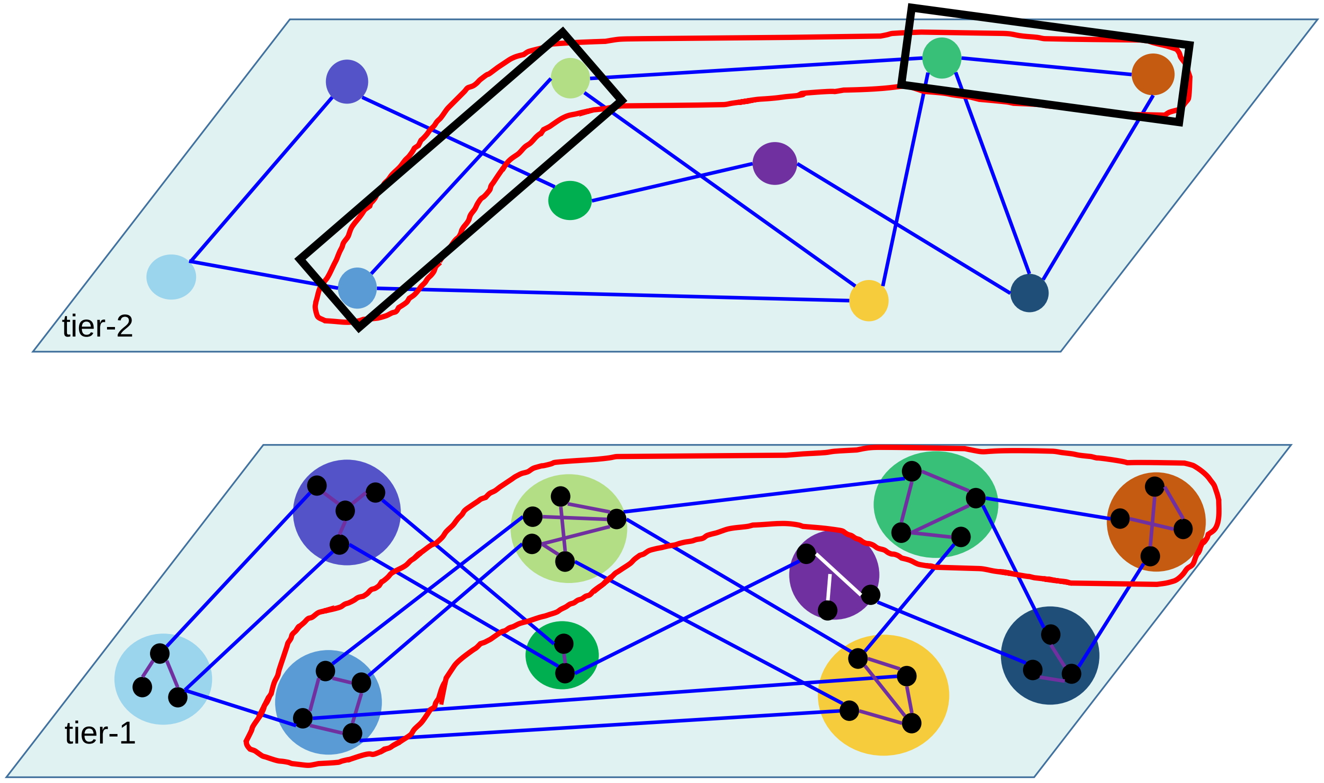
U: RV of distance between two random nodes in the first and last domains of k domains connected in a bus topology, with parameter $\beta = 1$

$F_U(d)$: CDF of RV U

Inter-domain routing strategies - partial sync among domains

Tier-2 topology: select domain-wise route ➡ **Form "synced units" on the domain-wise route** ➡ **Sync based on these "synced units"**

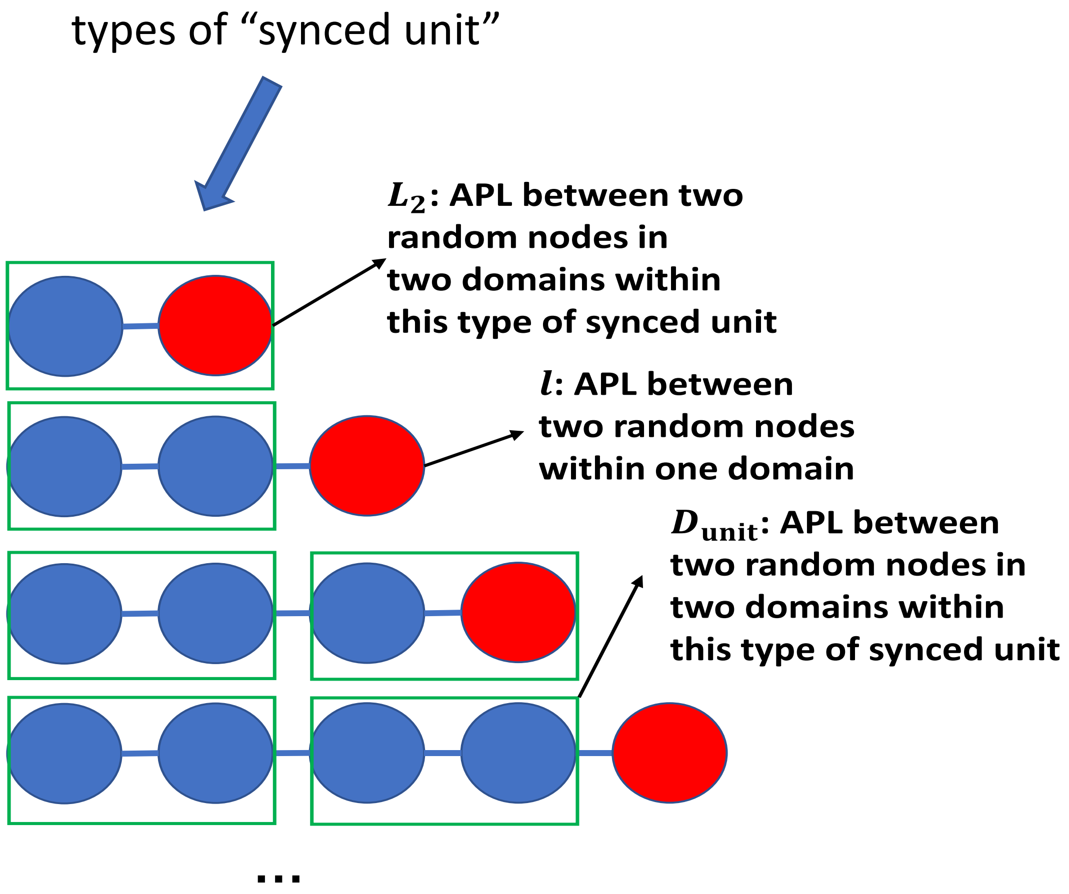
- **(1) The domain-wise path is jointly constructed by each controller in these domains, like BGP;**
- **(2) The SDN controller in current domain follows the instruction from the previous domain(s); if no such instruction exists, go to (3);**
- **(3) The SDN controller in the current domain selects a path starting from the ingress node to the closest egress node, and passes on the route selection results to the next domain**



Note: We don't intend to propose any SND controller synchronization protocol.

This section is only to demonstrate how to apply of our proposed analytical model in analysis

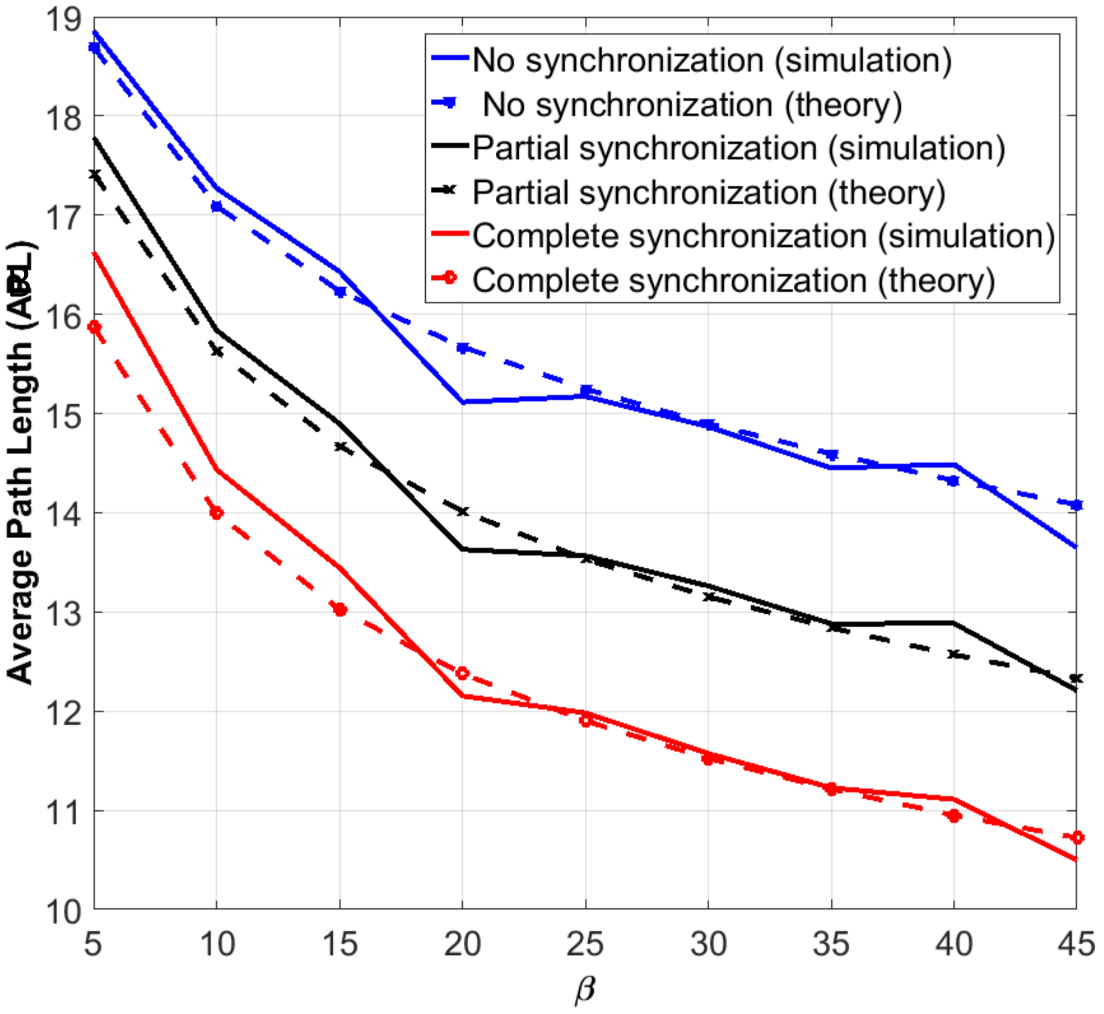
Main analytical results - partial sync among domains



- $L_k^{\text{SDN}} = \begin{cases} (\frac{k}{2} - 1)L_{\text{unit}} + L_2 + \frac{k}{2} - 1 & k \text{ is even,} \\ \frac{k-1}{2} L_{\text{unit}} + l + \frac{k-1}{2} & k \text{ is odd.} \end{cases}$
- $L_{\text{SDN}} = \sum_{y=2}^{m} L_y^{\text{SDN}} h_Y(y)$

L_k^{SDN}: APL in a bus topology with k domains
L_{unit}: APL in step (3)
L_{SDN}: APL under the simple scheme

Evaluations and results



- Simulation settings:
  - $n = 350$; (# nodes each domain)
  - $m = 50$; (# domain)
  - Intra-domain degree distribution collected from RocketFuel Project
  - Inter- domain degree distribution synthesised due to the lack of actual dataset

- Simulation results confirm the accuracy of our analytical framework
- With a given and limited synchronization level, the gap to optimal value by around can be reduced by 50%

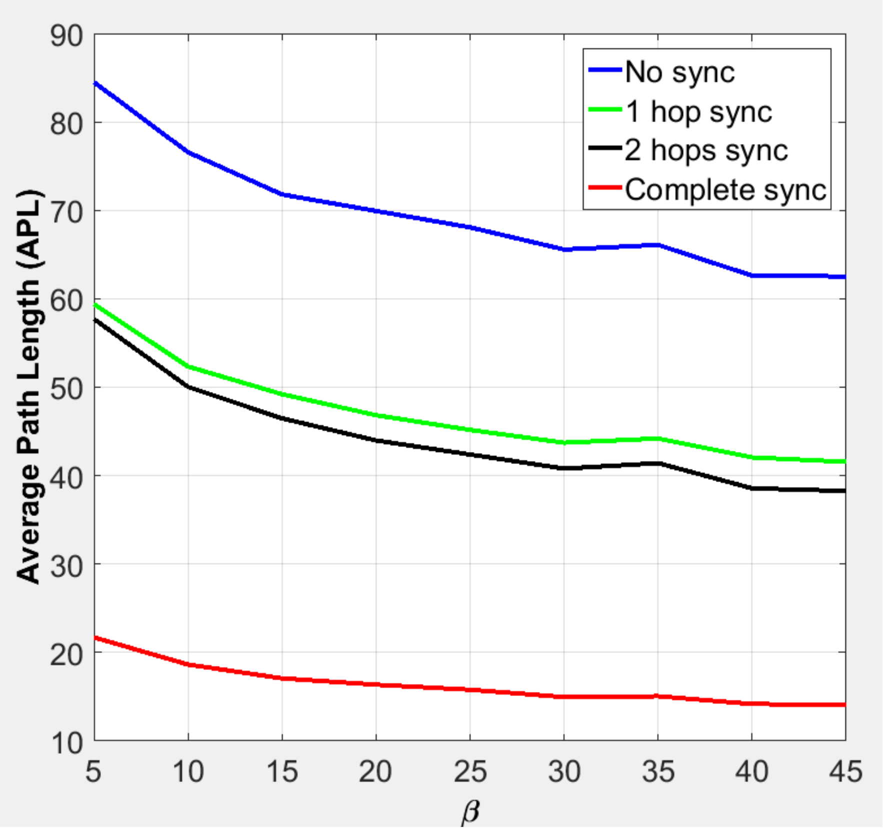
Summary of our work

- Goal: gain fundamental (analytical) understandings of performance enhancement offered by SDN

- Main contributions:
  - a generic 2-tier network model
  - analytical framework quantifying performance enhancement under different level of SDN controller synchronizations
  - simulation results confirming the accuracy of our developed analytical framework

Future plans

- Map network settings/status to edge weight in the network graphs of each domain
- Extend current analytical framework to heterogeneous edge weight scenarios
- Consider relaxing more system assumptions, such as the inter-domain connection parameter β
- Take into consideration some SDN controller placement-related problems in extending current work (brain storm sessions and discussions with colleagues)



Non-uniform edge weights scenarios: Initial simulation results

Questions?

Thank you !