Lyapunov Stability Analysis of Load Balancing in Datacenter Networks

Amrith Dhananjayan[1], Kiam Tian Seow[1] and Chuan Heng Foh[2]

[1]School of Computer Engineering, Nanyang Technological University, Republic of Singapore 639798.
[2]Department of Electronic Engineering, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom.

***Abstract*—Modern datacenters are becoming increasingly complex, with datacenter networks (DCN's) built to meet the data transmission demands of densely interconnected nodes of server hosts and switches. Load balancing in DCN's - to balance the bandwidth utilization among the DCN links - is indispensable for network stability as well as for meeting important objectives such as maximizing throughput and minimizing latency. Simulation has been the de facto empirical method for investigating the stability of DCN's under load balancing policies. To complement simulation with analytical insights into load balancing stability of DCN's, in this paper, we present an application of the qualitative version of the Lyapunov stability theory for load balancing DCN's modeled as discrete-event systems. The general Lyapunov control theory states that a controlled system is stable if there exists a function on the state space of the system, called the Lyapunov function, whose value is non-increasing along any execution trajectory of the system. Analytically proving the existence of such a Lyapunov function is sufficient to verify that the DES model representing a class of DCN's under a load balancing policy is stable. We illustrate the utility of our approach by investigating the stability of a class of DCN's configured in a fat-tree topology under a specific load balancing policy. Our work represents the first step towards a general qualitative stability theory for the policy design of load balancing algorithms for DCN's.**

***Index Terms*—Datacenter, discrete-event systems, Lyapunov stability**

I. INTRODUCTION

The increasing popularity of cloud-based services has driven the emergence of datacenters that have hundreds of thousands of server hosts and switches, with the numbers increasing exponentially fast [1]. These large and complex datacenters have built-in networks to satisfy the data transmission demands of densely interconnected nodes of hosts and switches. By spreading data transmission along multiple paths, load balancing in datacenter networks (DCN's) helps to maximize system throughput, minimize network latency and reduce congestion, which are the key objectives of DCN's [2]. Inefficient load balancing policies, however, can fail to rebalance the load distribution and cause some links to be congested while the others are underutilized, either permanently or for a prolonged period. This can result in network instability, with considerable reduction in network throughput as noted in [3].

This work was supported by the Singapore Agency for Science, Technology and Research, under SERC Future Data Center Technologies Programme Grant No: 112-172-0012. The authors can be contacted at amri0005@e.ntu.edu.sg, asktseow@ntu.edu.sg and c.foh@surrey.ac.uk.

The current state-of-the-art approach for investigating load balancing policies involves simulation studies as the de facto technique. However extensive, simulation is often practically performed in a non-exhaustive manner, checking policy-driven system stability only for selected test cases, and this presents a strong case for using qualitative analysis tools to theoretically prove system stability and complement simulation. Such analyses can give clear insights into the general system behavior of load balancing policies. However, to the best of our knowledge, there is no reported work in the literature on qualitative stability analysis that supports the policy design of load balancing algorithms for DCN's.

It has long been known that stability analysis can be done in a very broad setting described in terms of system dynamics, and which does not require explicit formulation of the system under investigation in terms of specific equations, differential or otherwise [4]. The theory of Lyapunov stability [5], which is due to the Russian mathematician Aleksandr Mikhailovich Lyapunov, is the most widely used control-theoretic tool for investigating the stability of dynamical systems. This versatile tool allows a system designer to prove the stability of a system if some function on the state space of the system, called the Lyapunov function, can be found whose value is non-increasing along any execution trajectory of the system. In essence, the existence of such a Lyapunov function is sufficient to verify the stability of a system.

In this paper, we take a novel qualitative approach to the stability analysis of DCN's under load balancing policies. DCN's are an important source of discrete-event systems (DES's) [6] - a formalism in which a system has discrete states, and for which state changes are driven by qualitative occurrences called events. We model (the load balancing operations of) DCN's as DES's [6] and utilize the Lyapunov stability theory developed for DES's [7] to analyze the modeled systems. In considering and modeling a class of DCN's configured in a fat-tree topology [8] under a specific load balancing policy, we identify a Lyapunov function and analytically prove the stability of this class of DCN's. In addition, we demonstrate that, in using the DES stability analysis approach, many useful insights into the qualitative conditions essential for load balancing stability can be uncovered. As the DES theory provides a general framework for the modeling and control of dynamical and highly complex systems [6], we believe that the theoretical research presented in this paper - on the discrete-event modeling and stability analysis of DCN's - constitutes the first step towards a general qualitative stability theory for the policy design of DCN's.

Our analysis assumes access to a global view of the available network resources and requests. This can be achieved in a DCN by setting up a software defined networking (SDN) environment [9] where switches are connected to a central controller running the OpenFlow protocol [10]. With OpenFlow being adopted in many modern DCN designs [8], [11], [12], [13] as multiple OpenFlow devices become more readily available in the market [14], [15], [16], our assumption is not unrealistic for stability analysis of modern DCN's.

On related work, there has been a lot of recent efforts that apply control theory for analyzing and improving datacenter environments. For example, in [17], Lyapunov optimization [18] is used to design a resource and power management system for a virtualized datacenter. In [19], a control framework to reduce energy consumption while satisfying service-level agreement constraints is presented. In [20], load balancing is modeled as a feedback controller design problem, and is solved using multiple input multiple output linear control. In [21], a unified framework is presented in which the dynamic behavior of a datacenter is modeled by a group of state-space models, and a set-theoretic control technique is used to solve the resource allocation issues identified as a common decision problem. There is, however, relatively little or no theoretical work that applies the qualitative theory of DES Lyapunov stability to complement the experimental and simulation studies on load balancing for DCN's, as proposed in this paper.

The rest of the paper is organized as follows. Section II briefly reviews the relevant background in DES modeling and the Lyapunov stability theory for DES's. Section III presents the general modeling of DCN's for load balancing. Section IV analyzes the Lyapunov stability of a DES model for fat-tree DCN's under a specific load balancing policy. Section V discusses the implications of our analysis. Finally, Section VI concludes the paper.

II. BACKGROUND

In this section, we review the general modeling and stability framework of DES's as proposed in [7].

A. Discrete-Event System Model

A DES G is modeled as a 5-tuple

$$G = (\chi, \mathcal{E}, f_e, g, E_v),$$

where χ is the set of states and $\mathcal{E}$ is the set of events. State transitions are defined by $f_e : \chi \to \chi$, where $e \in \mathcal{E}$. An event $e \in \mathcal{E}$ may occur at state $\mathbf{x} \in \chi$ if it is the set defined by the enable function $g : \chi \to 2^{\mathcal{E}}$. The set of all event trajectories is denoted by E and $E_v \subseteq E$ is the set of valid event trajectories specified for the system. E_v represents the set of event trajectories that are physically possible in G. The set of all valid event trajectories that start from state $\mathbf{x_0} \in \chi$ is denoted by $E_v(\mathbf{x_0})$. For event sequence $E_k = e_0, e_1, \cdots, e_{k-1}$, let $X(\mathbf{x_0}, E_k, k)$ denote the state at position k of event sequence E_k starting from $\mathbf{x_0} \in \chi$.

A set $\chi_m \subset \chi$ is called invariant with respect to G if $X(\mathbf{x_0}, E_k, k) \in \chi_m$ for all $k > 0$ and all E_k such that $E_k E \in E_v(\mathbf{x_0})$ for some infinite event sequence E. Let $E_a \subseteq E_v$ be the set of valid event trajectories that are allowed in the DES.

B. Lyapunov Stability of DES's

Let $\rho : \chi \times \chi \to \mathbb{R}^+$ denote a *metric* on χ. For $\chi_z \subseteq \chi$, $\rho(\mathbf{x}, \chi_z) = \inf\{\rho(\mathbf{x}, \mathbf{x}') : \mathbf{x}' \in \chi_z\}$ denotes the distance from $\mathbf{x}$ to the set χ_z. For an arbitrary set $\chi_z \subset \chi$, its r-*neighborhood* is denoted by the set

$$S(\chi_z; r) = \{\mathbf{x} : 0 < \rho(\mathbf{x}, \chi_z) < r\},$$

where $r > 0$.

A continuous function $\psi : [0, r] \to \mathbb{R}^+$ is said to belong to class K (i.e., $\psi \in K$) if $\psi(0) = 0$ and ψ is strictly increasing on $[0, r]$.

If an arbitrary execution of a DES starts from a state near but outside an invariant (state) set, and stays so infinitely, then the DES is said to be stable (with respect to the invariant set). Alternatively, we say that the invariant set of the DES is stable. According to the Lyapunov stability theorem for DES's, a DES is stable if there exists a functional called the Lyapunov function whose value is non-increasing for all possible trajectories of the DES.

Theorem 1: For an invariant set $\chi_m \subset \chi$ of G to be stable in the sense of Lyapunov with respect to E_a, it is sufficient that in a neighborhood $S(\chi_m; r)$ of the set χ_m, there exist a specified functional V (called "Lyapunov function") and $\psi_1, \psi_2 \in K$ such that the following properties hold:

1) $\psi_1(\rho(\mathbf{x}, \chi_m)) \leq V(\mathbf{x}) \leq \psi_2(\rho(\mathbf{x}, \chi_m))$, and
2) $V(X(\mathbf{x_0}, E_k, k))$ is a non-increasing function for $\mathbf{x_0} \in S(\chi_m; r)$, for all E_k such that $E_k E \in E_a(\mathbf{x_0})$ and all $k \geq 0$.

In addition to Lyapunov stability, if the DES converges into the invariant set infinitely often, i.e., the DES will reenter the invariant set whenever it is perturbed out of the set, then the DES is said to be asymptotically stable.

Theorem 2: For an invariant set $\chi_m \subset \chi$ of G to be asymptotically stable in the sense of Lyapunov with respect to E_a, it is sufficient that in a sufficiently small neighborhood $S(\chi_m; r)$, there exist a specified functional V and $\psi_1, \psi_2 \in K$ that have the properties of Theorem 1 and the property that $V(X(\mathbf{x_0}, E_k, k)) \to 0$ as $k \to \infty$ for all E_k, such that $E_k E \in E_a(\mathbf{x_0})$ for all $k \geq 0$ as long as $X(\mathbf{x_0}, E_k, k) \in S(\chi_m; r)$.

III. MODELING OF DCN'S FOR LOAD BALANCING

The physical structure of a DCN can be described by a directed graph, (H, L), where the nodes in H are the hosts and switches, and the edges in $L \subset H \times H$ are the links connecting the switches and hosts or other switches.

Let the maximum number of concurrent flows in the DCN be a_{max}. Each flow $F_a, 0 < a \leq a_{max}$, is a 3-tuple

$F_a = (s_a, d_a, b_a)$, where the nodes s_a and d_a are the source and destination hosts, respectively and b_a is the required bandwidth (in bits per second). We use $f_a(h_i, h_j) \in \{0, 1\}$ to denote the routing of flow F_a through the link (denoted by the edge) $(h_i, h_j) \in L$. It is 1 if the link (h_i, h_j) is on the route of flow F_a, and 0 otherwise.

Let $N = |L|$ be the number of links in the DCN and $\mathbf{Q}$ be the set of non-negative rational numbers. In what follows, for brevity, unless otherwise mentioned, we use a positive integer in $[1, N]$ to uniquely identify a link. For each link (h_i, h_j) identified as link n, $0 < n \leq N$, let $s(n)$ and $d(n)$ denote its source h_i and destination h_j, respectively. It then follows that the load on a link n is given by $x_n = \sum_{a=1}^{a_{max}} f_a(n) b_a$ (bits per second). At any arbitrary time k, all x_n's together constitute a DCN state $\mathbf{x_k}$, which provides a "timed snapshot" of the load distribution among the links in the DCN. Also, note that the load on a link cannot exceed its capacity.

In essence, load balancing is to ensure that a DCN enters a balanced state (within a specified tolerance), i.e., within the limits of the specified tolerance, all its links are distributed or redistributed with the same amount of load. Let M be the largest bandwidth required by any of the concurrent flows. Then an important problem of interest is to ensure that load balancing the DCN is asymptotically stable in the sense of Lyapunov, i.e., a balanced state within M can always be entered or reentered.

In our analytical study on load balancing, we make the standard assumption that the rate of arrival or departure of flows is significantly lower than the rate at which load is balanced. We also assume that the total required load is always less than the total link capacity of the DCN.

IV. Load Balancing Fat-Tree Datacenter Networks: A Stability Study

Fat-tree topology, which has received a lot of attention in DCN research, provides multiple paths among hosts. It is organized as a 3-layered hierarchical tree having switches at the core, aggregation and top-of-rack (ToR) layers as shown in Fig. 1. The hosts are connected to the switches in the ToR layer. The multipath support allows data flows to be distributed among various network components. Load balancing the traffic in these fat-tree networks is a problem of practical significance and is investigated in recent work such as [22], [23].

In this section, we present a DES model for fat-tree DCN's under a worst-fit load balancing policy, and examine their stability. With a fat-tree topology providing multiple links for the same flow, the worst-fit policy will select the link with the largest amount of available bandwidth. This policy can be easily implemented in a SDN controller [9]. Using the stability analysis technique reviewed in Section II, we shall prove that this class of DCN's modeled under worst-fit load balancing is stable in the sense of Lyapunov and asymptotically. Although our focus is on fat-tree topology which provides multipath support, our stability analysis result may be applicable to other topologies for which the constraints of the DCN model are satisfied.

A. DES Modeling

To begin with, let us formulate the worst-fit load-balancing flow model for fat-tree DCN's as a DES G. Let $\chi = \mathbf{Q}^N$ be the set of states, and $\mathbf{x_k} = [x_1, x_2, \cdots, x_N]^\intercal$ and $\mathbf{x_{k+1}} = [x'_1, x'_2, \cdots, x'_N]^\intercal$ denote the states at time k and $k+1$, respectively, where x_i is the load on link i. Let $x_i(k')$ denote the amount of load on link i at time k'. Let $e^{i,p(i)}_{\alpha(i)}$ represent the passage of load from link i to other links $m \in p(i)$, where $p(i) = \{j : j \neq i \text{ and } s(j) = s(i)\}$ denotes the alternative links to link i in the multipath fat-tree topology. Let the list $\alpha(i) = (\alpha_j(i), \alpha_{j'}(i), \cdots, \alpha_{j''}(i))$ such that $j, j', \cdots, j'' \in p(i)$, $j < j' < \cdots < j''$ and $\alpha_j(i) \geq 0$ for all $j \in p(i)$. The size of the list $\alpha(i)$ is $|p(i)|$. For convenience, the list is denoted by $\alpha(i) = (\alpha_j(i) : j \in p(i))$. For a link i, $\alpha_m(i)$ denotes the amount of load transferred from i to $m \in p(i)$. Let $\left\{e^{i,p(i)}_{\alpha(i)}\right\}$ denote the set of all such possible load transfers from link i. Then the set of events $\mathcal{E}$ of G is given by

$$\mathcal{E} = 2^{\bigcup_{i=1}^{N}\left\{e^{i,p(i)}_{\alpha(i)}\right\}} - \{\emptyset\}.$$

Now we specify g and f_e for $e_k \in g(\mathbf{x_k})$.

- Event $e^{i,p(i)}_{\alpha(i)} \in e_k$ such that $\alpha(i) = (\alpha_j(i) : j \in p(i))$ if both Conditions (a) and (b) specified below hold.
  - (a) i) $\alpha_j(i) = 0$, if $x_i - x_j \leq M$ where $j \in p(i)$,
    ii) $x_i - \sum_{m \in p(i)} \alpha_m(i) > \min_j\{x_j : j \in p(i)\}$,
    iii) $\alpha_{j^*}(i) > 0$, for some $j^* \in \{j : x_j \leq x_m \text{ for all } m \in p(i)\}$.
    Statement (i) of Condition (a) prevents the transfer of load from link i to link j if the loads on i and j are balanced within M. Statement (ii) requires that once some load is transferred from link i to another, the remaining load on link i must be larger than the load at time k in some alternative link $j \in p(i)$, to which that load could have been transferred. Statement (iii) requires that if link i is not balanced within M with all other links to which the load on link i could have been transferred (i.e., links in $p(i)$), then some load must be transferred from i to the least loaded link j^* in $p(i)$.
  - (b) $e^{i,p(i)}_{\alpha'(i)} \notin e_k$ where $\alpha'(i) = \{\alpha'_j(i) : j \in p(i)\}$ if $\alpha_j(i) = \alpha'_j(i)$ for some $j \in p(i)$ and $e^{i,p(i)}_{\alpha(i)} \in e_k$. Condition (b) ensures that the load transferred from link i to link j is consistently defined, in that, at time k, only some well-defined load is transferred.
- If $e_k \in g(\mathbf{x_k})$ and $e^{i,p(i)}_{\alpha(i)} \in e_k$, then $f_{e_k}(\mathbf{x_k}) = \mathbf{x_{k+1}}$, where

$$x'_i = x_i + \sum_{\{j : i \in p(j), e^{j,p(j)}_{\alpha(j)} \in e_k\}} \alpha_i(j) - \sum_{\{j : j \in p(i)\}} \alpha_j(i).$$

  The load x'_i on link i at time $k+1$ is the load on link i at time k plus the total load received by link i at time

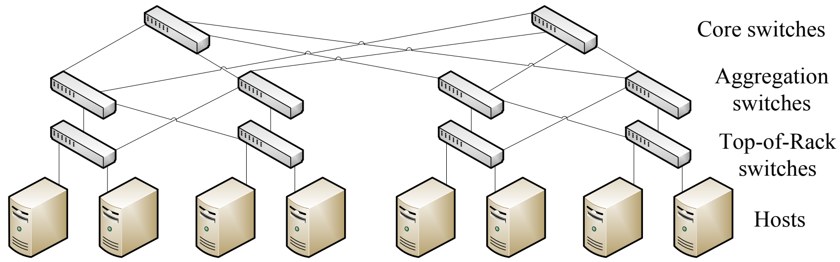

Fig. 1. DCN in a fat-tree topology

k minus the total load removed from link i at time k.

Let E_v be the set of all event trajectories that can be generated by the DES model defined by χ, $\mathcal{E}$, f_e and g. To specify E_i, we define an event $e^{i,i'}$ to represent the transfer of load from link i to some link $i' \in p(i)$. Let $E_i \subseteq E_v$ be the set of event trajectories in which events for each link can occur infinitely often on each $E \in E_i$.

B. Lyapunov Stability Analysis Result: Asymptotic Convergence to an M Balanced State

Consider the state set χ_b given by

$$\chi_b = \{\mathbf{x_k} \in \chi : |x_i - x_j| \leq M, \forall x_i, x_j \in \mathbf{x_k}, j \in p(i)\}. \tag{1}$$

χ_b is clearly an invariant set representing all the (load distribution) states in the DCN G that are balanced within M.

In what follows, we apply the Lyapunov stability theory [7] reviewed and show that the model G, representing fat-tree DCN's under a worst-fit load balancing policy, is asymptotically stable in the sense of Lyapunov.

Recall that $\mathbf{x_k} = [x_1, x_2, \cdots, x_N]^\intercal$. Let $\overline{\mathbf{x}} = [\overline{x}_1, \overline{x}_2, \cdots, \overline{x}_N]^\intercal$ be an arbitrary state in set χ_b and choose

$$\rho(\mathbf{x_k}, \chi_b) = \inf\{\max\{|x_i - \overline{x}_i| : i \in (0, N]\} : \overline{\mathbf{x}} \in \chi_b\}. \tag{2}$$

The main contribution of this paper may now be stated as follows.

Theorem 3: For the DCN G, the invariant set χ_b is asymptotically stable in the large with respect to E_i.

Proof: Choose

$$V(\mathbf{x_k}) = \begin{cases} \frac{1}{N}\sum\limits_{i=1}^{N} x_i - \min_i\{x_i\} & , \mathbf{x_k} \notin \chi_b \\ 0 & , \mathbf{x_k} \in \chi_b \end{cases}. \tag{3}$$

We now show that the selected Lyapunov function (3) satisfies Property 1 of Theorem 1.

According to Equation (3),

$$\rho(\mathbf{x_k}, \chi_b) \leq \max_i\{x_i\} - \min_i\{x_i\}. \tag{4}$$

As

$$\frac{1}{N}\sum_{i=i}^{N} x_i \geq \frac{1}{N}[\max_i\{x_i\} + (N-1)\min_i\{x_i\}],$$

from Equation (3) we have

$$V(\mathbf{x_k}) \geq \frac{1}{N}[\max_i\{x_i\} + (N-1)\min_i\{x_i\}] - \min_i\{x_i\},$$

i.e.,

$$V(\mathbf{x_k}) \geq \frac{1}{N}[\max_i\{x_i\} - \min_i\{x_i\}]$$

It follows from Equation (4) that

$$\frac{1}{N}\rho(\mathbf{x_k}, \chi_b) \leq V(\mathbf{x_k}). \tag{5}$$

According to Equation (1), for $\mathbf{x_k} \notin \chi_b$, there must be two links i and j, $j \in p(i)$, such that $x_i - x_j > M$. For the DCN to be balanced it is required that $x_i - x_j \leq M$. It follows that

$$\rho(\mathbf{x_k}, \chi_b) \geq \frac{1}{2}\max\{x_i - x_j - M : j \in p(i)\}.$$

Let $\pi_1(\mathbf{x_k}) = \max\{x_i - x_j - M : j \in p(i)\}$. Then

$$2\rho(\mathbf{x_k}, \chi_b) \geq \pi_1(\mathbf{x_k}). \tag{6}$$

As $\max_i\{x_i\} \geq \frac{1}{N}\sum\limits_{i=1}^{N} x_i$, according to Equation (3), $V(\mathbf{x_k}) \leq \max_i\{x_i\} - \min_i\{x_i\}$.

Because all N links are interconnected, it must be true that $\max_i\{x_i\} - \min_i\{x_i\} \leq N \max\{x_i - x_j, j \in p(i)\}$. Let $\pi_2(\mathbf{x_k}) = \max\{x_i - x_j, j \in p(i)\}$. Therefore,

$$V(\mathbf{x_k}) \leq N\pi_2(\mathbf{x_k}), \text{ i.e.,}$$

$$\frac{1}{N}V(\mathbf{x_k}) \leq \pi_2(\mathbf{x_k}). \tag{7}$$

From definitions of π_1 and π_2 it is clear that

$$\pi_1(\mathbf{x_k}) + M = \pi_2(\mathbf{x_k}). \tag{8}$$

Let us find a constant b such that $b\pi_1(\mathbf{x_k}) \geq \pi_2(\mathbf{x_k})$. For this we need $b \geq \frac{M}{\pi_1(\mathbf{x_k})} + 1$. It follows that, for $\mathbf{x_k} \notin \chi_b$, $b \in (1, \infty)$ is a real number and by Equations (6) and (7) that

$$2b\rho(\mathbf{x_k}, \chi_b) \geq b\pi_1(\mathbf{x_k}) \geq \pi_2(\mathbf{x_k}) \geq \frac{1}{N}V(\mathbf{x_k}).$$

As $\rho(\mathbf{x_k}, \chi_b) = V(\mathbf{x_k}) = 0$ for $\mathbf{x_k} \in \chi_b$, we have

$$2bN\rho(\mathbf{x_k}, \chi_b) \leq V(\mathbf{x_k}) \text{ for all } \mathbf{x_k} \in \chi. \tag{9}$$

From Equations (5) and (9), it follows that

$$\frac{1}{N}\rho(\mathbf{x_k}, \chi_b) \leq V(\mathbf{x_k}) \leq 2bN\rho(\mathbf{x_k}, \chi_b),$$

satisfying Property 1 of Theorem 1.

We now show that the Lyapunov function (3) satisfies Property 2 of Theorem 1.

According to Equation (3), $V(\mathbf{x_k})$ varies only with the lightest load $\min_i\{x_i\}$ of the DCN. The policy prevents the transfer of load on this least loaded link to another link. Also, transfer of load $\alpha(l)$ from some link l to the least loaded link is possible only if $x_l > \min_i\{x_i\} + \alpha(l)$. Thus $V(\mathbf{x_{k+1}}) \geq V(\mathbf{x_k})$, proving Property 2 of Theorem 1.

Continuing from these arguments, we now prove that χ_b is asymptotically stable in the large with respect to E_i. The policy always attempts to transfer load to the least loaded link and, as a result, the value of $\min_i\{x_i\}$ increases over time. Therefore, for every $k \geq 0$, there exists $k' > k$ such that $V(\mathbf{x_{k'}}) < V(\mathbf{x_k})$, as long as $\mathbf{x_{k'}} \notin \chi_b$. So, as $k \to \infty$, $V(X(\mathbf{x_0}, E_k, k)) \to 0$, proving the extra property stated in Theorem 2 for asymptotic stability.

This completes the proof. ∎

V. DISCUSSION ON LYAPUNOV STABILITY ANALYSIS

Our stability analysis shows that the fat-tree DCN model G containing the worst-fit policy of transferring flow to the least loaded link is asymptotically stable in the sense of Lyapunov. It means that, under the worst-fit policy, all the links in a fat-tree DCN would be distributed or redistributed with the same amount of load (within tolerance M).

Alternative load assignment policies for fat-tree DCN's, such as first-fit [8], [13] and best-fit [22], have also been proposed. The first-fit policy simply selects, among all possible links, the first one it finds with sufficient bandwidth; and the best-fit policy selects a link with the least remaining bandwidth that is sufficient for the load. These two policies have a tendency to accumulate load over a small number of links and are therefore prone to congestion [22]. The conjecture is that a DES model for fat-tree DCN's under either of these two policies is not asymptotically stable, and therefore no Lyapunov function exists that satisfies all the properties stated in Theorem 2. However, further studies are needed to confirm this conjecture.

It has to be noted that establishing the existence of a suitable Lyapunov function is sufficient but not necessary to prove the stability of DCN's with respect to a load balancing policy. This means that failing to identify a Lyapunov function does not in general suggest that a DCN model is unstable.

Using our reported study as a stepping stone, future work can delve into qualitative stability analysis of different policy-driven DCN's by applying our proposed approach, which furnishes a useful tool that complements simulation.

VI. CONCLUSION

The qualitative version of the theory of Lyapunov stability is proposed for analytically verifying and providing insights into the stability of load balancing policies in DCN's. Our approach complements simulation, the de facto method for investigating the DCN stability of load balancing policies. The non-exhaustive manner with which simulation is performed, where stability is checked only for selected test cases, presents a strong case for our work. We have illustrated the utility of our approach by investigating the stability of a class of DCN's under a worst-fit load balancing policy, configured in a fat-tree topology. Future work includes extending the study to any routing protocol that poses stability uncertainty.

REFERENCES

[1] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, "Dcell: A scalable and fault-tolerant network structure for data centers," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, 2008, pp. 75–86.

[2] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, 2008, pp. 63–74.

[3] D. Abts and B. Felderman, "A guided tour through data-center networking," *Queue*, vol. 10, no. 5, p. 10, 2012.

[4] K. M. Passino, A. N. Michel, and P. J. Antsaklis, "Lyapunov stability of a class of discrete event systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 2, pp. 269 – 279, 1994.

[5] V. I. Zubov and L. F. Boron, *Methods of AM Lyapunov and their Application*. Noordhoff Groningen, 1964.

[6] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Springer, 2008.

[7] K. M. Passino and K. L. Burgess, *Stability Analysis of Discrete Event Systems*. Wiley New York, 1998.

[8] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, "Hedera: Dynamic flow scheduling for data center networks," in *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, 2010, pp. 19–19.

[9] O. M. E. Committee *et al.*, "Software-defined networking: The new norm for networks," *ONF White Paper. Palo Alto, US: Open Networking Foundation*, 2012.

[10] "The OpenFlow Consortium," http://www.openflowswitch.org.

[11] T. Benson, A. Anand, A. Akella, and M. Zhang, "The case for fine-grained traffic engineering in data-centers," *INM/WREN*, vol. 10, 2010.

[12] R. N. Mysore, A. Pamboris, N. Farrington, N. Huang, P. Miri, S. Radhakrishnan, V. Subramanya, and A. Vahdat, "Portland: A scalable fault-tolerant layer 2 data center network fabric," in *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 4, 2009, pp. 39–50.

[13] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yiakoumis, P. Sharma, S. Banerjee, and N. McKeown, "Elastictree: Saving energy in data center networks," in *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, 2010, pp. 17–17.

[14] "NEC OpenFlow Switches," http://www.necam.com/pflow/.

[15] "OpenFlow 1.1 Release," www.openflowswitch.org/wk/index.php.

[16] "Cisco OpenFlow Switches," http://blogs.cisco.com/tag/openflow/.

[17] R. Urgaonkar, U. C. Kozat, K. Igarashi, and M. J. Neely, "Dynamic resource allocation and power management in virtualized data centers," in *Network Operations and Management Symposium (NOMS), 2010 IEEE*. IEEE, 2010, pp. 479–486.

[18] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource Allocation and Cross-layer Control in Wireless Networks*. Now Pub, 2006.

[19] D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy, and G. Jiang, "Power and performance management of virtualized computing environments via lookahead control," *Cluster computing*, vol. 12, no. 1, pp. 1–15, 2009.

[20] Y. Diao, J. L. Hellerstein, A. J. Storm, M. Surendra, S. Lightstone, S. Parekh, and C. Garcia-Arellano, "Using mimo linear control for load balancing in computing systems," in *Proceedings of the 2004 IEEE American Control Conference*, vol. 3, 2004, pp. 2045–2050.

[21] D. Dechouniotis, N. Leontiou, N. Athanasopoulos, G. Bitsoris, and S. Denazis, "Acra: A unified admission control and resource allocation framework for virtualized environments," in *8th International Conference on Network and Service Management (CNSM)*, 2012, pp. 145–149.

[22] T. Cheocherngngarn, H. Jin, J. Andrian, D. Pan, and J. Liu, "Depth-first worst-fit search based multipath routing for data center networks," in *IEEE Global Communications Conference (GLOBECOM)*, 2012, pp. 2821–2826.

[23] S. Mahapatra, X. Yuan, and W. Nienaber, "Limited multi-path routing on extended generalized fat-trees," in *IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW)*, 2012, pp. 938–945.