The effect of network topology on the control traffic in distributed SDN

Muhammad Zeshan Naseer and Viktoria Fodor
Department of Network and System Engineering, KTH - Royal Institute of Technology, Stockholm
{mznaseer, vjfodor}@kth.se

***Abstract*—Software Defined Networking (SDN) has the promise of flexible routing, traffic management and service provisioning in communication networks. To allow SDN based networks scale in size, the control architecture needs to be distributed, which in turn requires the introduction of controller to controller communication. This is needed to ensure that the distributed controllers have the same understanding about the underlaying network and can make consistent local decisions. In this paper we evaluate the volume of the emerging control traffic, considering a distributed controller architecture based on ONOS and OpenFlow. We show that the control traffic increases drastically with the number of controllers, as well as with the size of the underlaying network. We evaluate topologies forming regular and random graphs, and conclude that the type of the topology influences the traffic volume significantly, while the network density has less significant effect. We show that the control traffic is significant even if the number of controllers is selected such that the control traffic is minimized, and we argue that further optimization of ONOS is needed to trade off control traffic load and consistency in the network views.**

Keywords*—*SDN, ONOS, control, scalability

I. Introduction

Software Defined Networking (SDN) is becoming a generally accepted solution to provide the increased flexibility needed for service differentiation and resource efficiency in wired as well as in wireless networks [1] [2]. In an SDN, a centralized controller takes over the control from the individual switching nodes. The network operating system collects the information about the network, and helps the controller to make an abstract model of the network topology. This complete knowledge of the network helps the controller to dynamically provision the network resources, to apply the concepts of fairness and traffic shaping, as well as to provide complex routing policies for enhancing security, achieving quality of service differentiation or to support network function visualization.

The use of a single controller however raises reliability and scalability issues. First, a single controller SDN becomes a single point of failure, which is critical for network performance and reliability. Second, a single controller cannot support a large network due to limited memory and processing capabilities. Finally, the position of a single controller is also critical in terms of switch to controller network delays. As a result, a network based on a single controller would have scalability constraints based on memory, processing capability and reaction time.

For improved scalability of the network, a logically centralized controller can be implemented through a cluster of controllers, leading to a distributed SDN architecture. This architecture permits to balance the switch to controller traffic among different controllers, improves reliability and can achieve scalability by limiting the load of a single controller and the switch to controller network delays. However, this architecture also implies that the controllers must coordinate to obtain a consistent view of the network state [3]. As the set of controllers need to share information about the switches they own and about the network topology, controller to controller traffic is introduced which may be non-negligible as the number of controllers or the network itself grows [4].

The objective of this paper is to evaluate the effect of the network topology on the emerging control traffic, and to discuss how these traffic can be minimized. We estimate the volume of the emerging control traffic in large distributed SDNs, that applies ONOS for controller to controller and OpenFlow for switch to controller communication. We utilize the measurement results of [4], and build a model for the controller to controller traffic for general network topologies. Then we consider the specific cases of regular and random network topologies, and evaluate the effect of the network size and of the number of controllers. We show that for the same network size and density, the control traffic is higher in regular networks, but the scalability properties in the two topologies are similar. The optimal number of controllers depends significantly on the network size and on the intensity of the switch to controller queries, while the density of the network has only marginal effect.

II. Related work

The placement of controllers is one of the main issues in the design of distributed SDNs and the literature addresses controller placement for various objectives. The seminal paper [5] proposes controller placement heuristics that trade off latency, failure tolerance and load balancing. In [6], [7] a given set of controllers is placed, such that the computational capacity of the controllers is obeyed and the maximum switch to controller delay is minimized, while [8] minimizes the cost of controller deployment and maintenance, keeping the constraints on the controller and network capacities as well as

ISBN 978-3-903176-08-9 2018 IFIP

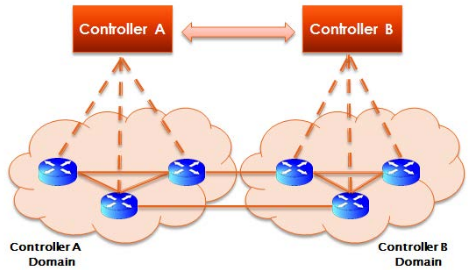

Fig. 1. Distributed SDN architecture. Controllers A and B exchange information to form a virtual centralized controller.

on the switch to controller delays. Reliable distributed SDN is considered in [9], [10], where delay and capacity constraints have to be fulfilled even when backup controllers are used.

There are only few works addressing the controller placement problem from the point of view of the control traffic, since network operating systems for distributed architectures are in their infancy. The emerging ONOS is evaluated in [4], by emulating small networks with the objective of finding mathematical models of the controller to controller traffic load. These findings are then used in [11] to evaluate the effect of controller placement on the control traffic in given, large topologies. In our work we build as well on the measurement results of [4] to systematically evaluate how the control traffic is affected by the network size and network topology. To estimate the rate of switch to controller traffic for flow setup, we use the measurement results of [12]. More exact models considering network topology as well as traffic matrix are also available in the literature [13], these could be considered for more detailed studies.

III. Distributed SDN

In this paper we consider a distributed software defined network with the general architecture shown on Figure 1. According to the SDN principles, the control functions are decoupled from the forwarding functions at the network switches, and are implemented by a set of controllers, each responsible for a cluster of switches. The controllers exchange information to have a common network view and to form a virtually centralized controller.

In this architecture two control planes can be identified: the switch to controller (s2c) plane , which is present already in a centralized SDN and supports the set up of the forwarding tables based on the transmission paths determined at the controller, and the controller to controller (c2c) plane, which ensures correct controller functionalities despite the distributed implementation. The availability of the shared data structures affects the process of providing information for the switches querying their controller. In this paper we consider the multiple data-ownership model, where each controller has a local copy of all data required for routing decisions, therefore switch queries never need to be forwarded in the c2c plane.

The s2c plane has rich literature with mature architecture designs and protocol implementations like OpenFlow [14]. The development of the c2c plane is less mature, where ONOS [15] and OpenDaylight [16] are two prominent open implementations. In this work we build our analysis on the characteristics of ONOS, but our results could be generalized for other solutions.

The design of the c2c plane needs to address two main issues: first, the consistency of the shared view of the network graph, which is required for the correct control of the forwarding plane; and second, the availability of the shared data structures, required for fast decision making. Consistency in general can be strong or eventual. Strong consistency means that all available information is identical at all the controllers, while eventual consistency means that the controllers may have temporally different views, but these will converge with time. In systems with communication delays strong consistency compromises all time data availability [17]. Therefore, ONOS combines strong and eventual consistency models, to trade-off between delay and consistency:

1) The controllers need to have an agreement all the time on the assignment of the switches. Therefore, cluster membership is shared under the strong consistency model, using the Raft protocol [18]. Raft messages are exchanged when the membership of a cluster has changed – that is, a new switch got connected or a switch has left.
2) Other information, including the network topology, is shared under the eventual consistency model, using the so called anti-entropy algorithm [19], implementing simple gossiping among all the controllers. In the anti-entropy algorithm, controllers send out update messages periodically to randomly selected controllers and in this way the information on the changes eventually propagates in the network of controllers. The consensus time for the fully connected network of C controllers is $O(\ln C)$ [20].

IV. Control traffic model

A. Methodology

Our objective is to build up an SDN control traffic model, including both controller to controller (c2c) and switch to controller (s2c) traffic. Since control messages are transmitted through the network links over multihop paths, the model need to reflect the length of the transmission paths across the network that the control messages travel. Therefore, we derive the control traffic load in two steps: first we express the bandwidth (or rate) of the generated c2c and s2c control traffic, B_{c2c} and B_{s2c} and then weight them by the length of the transmission paths the control messages travel, to get the total control traffic $T_T = T_{c2c} + T_{s2c}$.

Our objective is to evaluate the effect of the topology on the control traffic load. Therefore, we consider two complementary topologies, regular grids and Erdős-Rényi random graphs. These graph structures represent two extremes in the world of network models: regular grids have large average path length, but also high clustering coefficient. On the other end, ER graphs have logarithmic increase in path length, but low clustering coefficient. These graph metrics are important in the case of distributed SDNs, for efficient controller to controller and switch to controller communication respectively.

For simplicity, throughout the paper we assume that fraction and square root operations give integer values. If it does not hold, the results are approximate.

B. Control traffic bandwidth

We consider a network of S switches and C controllers, forming a clustered distributed SDN architecture, where each controller is responsible for a subset of the switches.

The ONOS control traffic bandwidth have been measured systematically for networks with two and three controllers in [4]. The measurements revealed that the control traffic bandwidth between two controllers depends linearly on the number of switches and edges in the clusters and on the number of edges between the clusters. Using these results, we can build up a general model for networks with C controllers, and the control traffic generated by any controller i to any other controller can be expressed as

$$B_i = b^0 + b^s S_i + b^l L_i + b^{s-} \sum_{j\neq i} S_j + b^{l-} \sum_{j \neq i} L_j + b^d \sum_{j\neq i} L_{ij} + b^e \sum_{j\neq i} \sum_{k \neq i,j} L_{jk}, \tag{1}$$

where S_i is the number of switches in cluster i, L_i is the number of internal links in the cluster, and L_{jk} is the number of links where the end nodes belong to cluster j and k respectively. Parameter b^0 is the *zero bandwidth*, representing the control traffic that is generated even when the network does not contain any switches. Parameters b^s and b^l represent control information about switches and links in the own cluster, b^{s-} and b^{l-} about switches and links within other clusters, b^d about links with one end in the own cluster and b^e on links running among other clusters.

The measurement results in [4] show that the b^* (the * may denote s, l, $s-$, $l-$, d or e) parameters depend on the number of controllers C in the network. Due to the anti-entropy protocol, where the controller to communicate to is selected randomly, we could expect that $b^* \propto 1/C$, but the measured values for $C = 2$ and 3 show that b^* includes also a fix part. Therefore, we consider

$$b^* = b^*_f + \frac{1}{C-1} b^*_p, \tag{2}$$

where b^*_f is the fix part, and b^*_p contributes to the part that scales down inversely proportionally with C.

Table I shows the parameter values estimated from the measurements results in [4]. Note that from (2) we have $b^d_f = 0$, therefore we estimate $b^e_f \approx 0$. From the table we can also conclude that the dominating parameter is the zero bandwidth b^0 for small networks. For large networks the switch and link related control traffic has similar contribution.

TABLE I
BANDWIDTH PARAMETER VALUES. THE UNIT IS KBPS, THE $C = 2$ AND $C = 3$ VALUES ARE FROM [4]

Parameter	$C = 2$	$C = 3$	b^*_f	b^*_p
b^0	53	43	33	20
b^s	1.45	1.12	0.79	0.66
b^l	1.45	0.93	0.41	1.04
b^{s-}	0.48	0.35	0.22	0.26
b^{l-}	0.94	0.53	0.12	0.82
b^d	1.55	0.87	*0*	*1.6*
b^e	-	0.7	*0*	*0.7*

Finally, we derive the aggregated generated c2c control traffic, including all traffic generated by all the controllers to all the other controllers, resulting

$$\begin{aligned} B_{c2c} &= \sum_{i=1}^{C} (C-1) B_i \\ &= C(b^0_f (C-1) + b^0_p) \\ &\quad + S(b^{s-}_f (C-1)^2 + b^{s-}_p (C-1) + b^s_f (C-1) + b^s_p) \\ &\quad + L(b^{l-}_f (C-1)^2 + b^{l-}_p (C-1) + b^l_f (C-1) + b^l_p) \\ &\quad + L_d (2b^d + (C-2) b^e), \end{aligned} \tag{3}$$

where L is the number of intra-cluster links, and L_d is the number of inter-cluster links. Note that (3) provides a closed form expression of the control traffic that does not depend on the characteristics of the single clusters. We can conclude that the control traffic depends on the total number of controllers C, switches S, intra-cluster links L, and inter-cluster links L_d. As we see, B depends linearly on S, L and L_d, but increases quadratically with C, which introduces significant penalty as the number of controllers is increased.

The controllers also communicate with the switches in their own cluster, e.g., through the OpenFlow protocol, where switches query the controllers for routing information. We express the aggregated switch to controller traffic as

$$B_{s2c} = (S - C) 2 \lambda G, \tag{4}$$

where λ is the long term average of the number of messages generated by a switch per second, G is the average message size in bits. The expression reflects that the C controllers are hosted by C switches, and these switches do not generate s2c control traffic on the network links.

To proceed and derive the actual volume of control traffic in the network, we need to consider the length of the paths the control traffic is forwarded on. Therefore, from this point we consider specific topologies.

C. Network control traffic in regular grid topology

First we consider a wrapped two dimensional grid topology. We present detailed results for the rectangular grid, where each node has $k = 4$ neighbors, as shown on Figure 2. Then we generalize the results for some specific $k > 4$ values to see how k affects the control load.

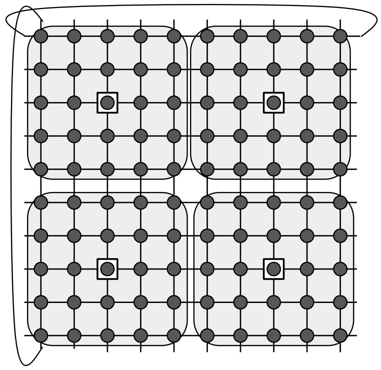

Fig. 2. Network forming a regular, folded grid with $k = 4$. Controllers (marked by squares) are placed regularly. Switches join the closest controller, which results equal size clusters (marked by shaded areas).

The S switches form C rectangular clusters of size $S_C = S/C$, and the controller is placed in one of the central switches to minimize the s2c control traffic as well as the maximum s2c delay. In the regular grid topology all controllers have statistically the same location, and consequently, all the B_i values are identical. Let P_{ij} be the length of the transmission path from controller i to controller j, and $\bar{P}^G_{c2c}$ the average controller to controller path length. We can then express the aggregate controller to controller load as

$$\begin{aligned} T^G_{c2c} &= \sum_{i=1}^{C} \sum_{j=1, j\neq i}^{C} B_i P_{ij} \\ &= B_i \sum_{i=1}^{C} \sum_{j=1, j\neq i}^{C} P_{ij} \\ &= B_i C(C-1) \frac{\sum_{i=1}^{C} \sum_{j=1, j\neq i}^{C} P_{ij}}{C(C-1)} \\ &= C(C-1) B_i \bar{P}^G_{c2c} \\ &= B^G_{c2c} \bar{P}^G_{c2c}. \end{aligned} \tag{5}$$

To express B^G_{c2c}, we need to replace the topology dependent parameters in (3), that is, the number of inter-cluster links and the number of intra-cluster links with the grid topology specific values. In the case of $k = 4$, the number of inter-cluster links leaving one cluster is equal to the length of the cluster border $4\sqrt{S_C}$, thus, for the entire network we get

$$L^G_d = C \frac{4\sqrt{S_C}}{2} = 2\sqrt{SC}, \tag{6}$$

then, sice altogether there are $2S$ links, the number of intra cluster links can be calculated as

$$L^G = 2S - 2\sqrt{SC}. \tag{7}$$

Finally, we need to derive the average path length $\bar{P}^G_{c2c}$. We notice that in the considered topology, shortest paths connecting far away controllers can be constructed from path segments between neighboring controllers. Therefore

$$\bar{P}^G_{c2c} = \bar{H}^G_{c2c} \sqrt{\frac{S}{C}}, \tag{8}$$

where $\sqrt{S/C}$ is the distance of neighboring controllers and $\bar{H}^G_{c2c}$ is the average number of path segments to travel from any controller to any controller.

Let us derive $\bar{H}^G_{c2c}$. We restrict the derivation for odd $\sqrt{C}$ values. We can see that each controller needs to reach other controllers in maximum $(\sqrt{C} - 1)/2$ steps away in both horizontal and vertical directions. Therefore, for odd $\sqrt{C}$ we get the average number of path segments to travel

$$\bar{H}^G_{c2c} = \frac{\sqrt{C}}{2}, \tag{9}$$

which provides

$$\bar{P}^G_{c2c} = \frac{\sqrt{C}}{2} \sqrt{\frac{S}{C}} = \frac{\sqrt{S}}{2}. \tag{10}$$

$\bar{H}_{c2c}$ is somewhat higher for even $\sqrt{C}$ values, but the difference diminishes as C increases, and therefore we use (10) in the followings.

Substituting (3) and (10) into (5), we get

$$\begin{aligned} T^G_{c2c} = \frac{S^{3/2}}{2} \Bigg[& \frac{1}{S_C} \Big(b^0_f \Big(\frac{S}{S_C} - 1 \Big) + b^0_p \Big) \\ & + \Big(b^{s-}_f \Big(\frac{S}{S_C} - 1 \Big)^2 + (b^{s-}_p + b^s_f) \Big(\frac{S}{S_C} - 1 \Big) + b^s_p \Big) \\ & + 2 \Big(1 - \frac{1}{\sqrt{S_C}} \Big) \Big(b^{l-}_f \Big(\frac{S}{S_C} - 1 \Big)^2 \\ & \qquad + (b^{l-}_p + b^l_f) \Big(\frac{S}{S_C} - 1 \Big) + b^l_p \Big) \\ & + \frac{2}{\sqrt{S_C}} \Big(2b^d + \Big(\frac{S}{S_C} - 2 \Big) b^e \Big) \Bigg], \end{aligned} \tag{11}$$

where $S_C = S/C$ is the size of the clusters. From (11) we can conclude that the controller to controller traffic increases with $S^{7/2}$, while it decreases inversely proportionally with the cluster size. This would motivate the use of few controllers, to decrease the controller to controller traffic.

Let us now consider the control traffic generated by the switch to controller messages, given by (4). The optimal position of the controller is in the middle of a cluster. Then the average length of the switch to controller transmission path can be calculated similarly to $\bar{H}^G_{c2c}$, but considering the grid of switches, instead of the grid of controllers, which gives $\bar{P}^G_{s2c} = \sqrt{S_C}/2$, and the total switch to controller traffic load becomes

$$T^G_{s2c} = B^G_{s2c} \bar{P}^G_{s2c} = \lambda G S \sqrt{S_C} \Big(1 - \frac{1}{S_C} \Big). \tag{12}$$

From (11) and (12) we see that T_{c2c} decreases, while T_{s2c} increases with the cluster size S_C. Therefore, as expected, there is a cluster size that minimizes the control traffic for

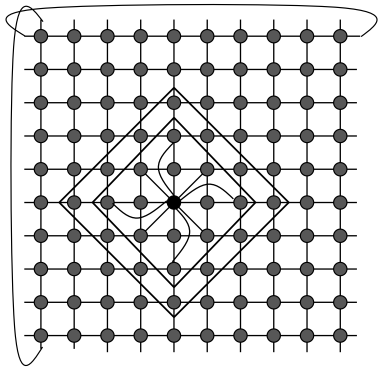

Fig. 3. Network forming a regular, folded grid. Neighborhood areas for $k = 12$ and 24, resulting in $l = 2$ and 3 respectively are shown. For the $k = 12$ case we also show the connection pattern.

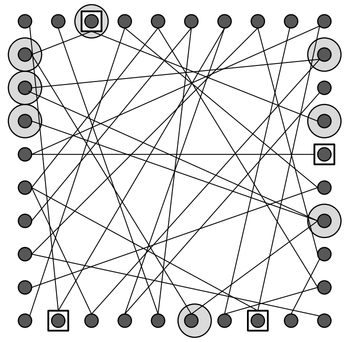

Fig. 4. Network forming an Erdős-Rényi random graph. Controllers (marked by squares) are placed randomly. Switches join the closest controller, which results clusters of variable size. Some nodes of the same cluster are marked by shaded circles.

given network size and the intensity of the switch to controller traffic λ.

To generalize the results, we consider grid topologies where each node can reach its k closest neighbors, resulting in a denser, but still regular grid. To simplify the derivations, while still addressing large networks, we consider the specific k values, where it holds that shortest c2c paths can be constructed using only horizontal and vertical edges. E.g., this is the case for $k = 12$ and 24, where single hop neighbors form a rectangular as shown on Figure 3. Also we assume that the cluster diameters are much larger than the longest single hop in the grid, and thus the derivation steps we followed for $k = 4$ are still valid. That is, we need to derive the k specific L^G, L^G_d, $\bar{P}^G_{c2c}$ and $\bar{P}^G_{s2c}$ expressions.

Let us denote by l the length of the longest transmission, measured in the distance of the neighboring nodes (e.g., for $k = 4$ $l = 1$, and for $k = 12$ shown on Figure 3, $l = 2$. Inspecting Figure 3 we can derive the relationship between k and l

$$k = 4(1+2+...+l) = 2l(l+1), \quad l = \frac{\sqrt{1+2k}-1}{2}. \quad (13)$$

We start by deriving L^G_d. Now all the nodes that are within distance l of the cluster border have inter-cluster links, giving

$$L^G_d = 2\sqrt{C}\sqrt{S}\sum_{i=0}^{l-1}(l-i)(2i+1) = \ldots = \frac{\sqrt{SC}k\sqrt{1+2k}}{6},$$

$$L^G = \frac{1}{2}kS - L^G_d. \quad (14)$$

Equation (9) still holds, while the distance of neighboring controllers becomes $\sqrt{S/C}/l$. This provides the average controller to controller transmission path

$$\bar{P}^G_{c2c} = \frac{\sqrt{S}}{2l} = \frac{\sqrt{S}}{\sqrt{1+2k}-1}. \quad (15)$$

We can see k has opposite effects on L^G, L^G_d and $\bar{P}^G_{c2c}$. The final T^G_{c2c} expression becomes

$$\begin{aligned} T^G_{c2c} = \frac{S^{3/2}}{\sqrt{1+2k}-1}\Bigg[\frac{1}{S_C}\Big(b^0_f\Big(\frac{S}{S_C}-1\Big)+b^0_p\Big) \quad (16)\\ +\Big(b^{s-}_f\Big(\frac{S}{S_C}-1\Big)^2+(b^{s-}_p+b^s_f)\Big(\frac{S}{S_C}-1\Big)+b^s_p\Big)\\ +\frac{k}{2}\Big(1-\frac{\sqrt{1+2k}}{3\sqrt{S_C}}\Big)\Big(b^{l-}_f\Big(\frac{S}{S_C}-1\Big)^2\\ +(b^{l-}_p+b^l_f)\Big(\frac{S}{S_C}-1\Big)+b^l_p\Big)\\ +\frac{k\sqrt{1+2k}}{6\sqrt{S_C}}\Big(2b^d+\Big(\frac{S}{S_C}-2\Big)b^e\Big)\Bigg]. \end{aligned}$$

The c2c traffic still increases with $S^{7/2}$. Larger part of the traffic decreases with $\sqrt{k}$, due to the longer single hop transmissions, however, the traffic has a component that increases linearly with k, due to the increasing number of edges that has to be reported in the ONOS messages.

Considering the switch to controller traffic, the longer links lead to shorter transmission paths, giving

$$\bar{P}^G_{s2c} = \frac{\sqrt{S_C}}{2l} = \frac{\sqrt{S_C}}{\sqrt{1+2k}-1}. \quad (17)$$

This makes T^G_{s2c} decrease with increasing $\sqrt{k}$, that is, the increased network density effects only positively the switch to controller traffic.

D. Network control traffic under random topology

Let us now consider a network with random topology, specifically, a network of S switches, and links forming an Erdős-Rényi (ER) random graph, where two switches are connected with probability p. Since all switches see statistically similar neighborhood, we place the C controllers at random switch locations, and we assume, that each of the switches connects to the closest controller. A possible realization of such a network is shown on Figure 4.

The ER graph is homogeneous in the sense that from each node the path length distribution to all other nodes can be characterized by the same CDF $F(x)$ and pdf $f(x)$. Then, the probability that a given switch i with path length x_i selects controller j as its closest controller is the probability that the paths to the other controllers are longer than x_i, that is $\bar{F}(x_i)^{C-1}$. The probability that an arbitrary node selects controller j becomes $p_j = \int \bar{F}(x)^{C-1} f(x)dx$. Note, that the right hand side of this equation is independent from j, and therefore has to be the same for each controller, that is, $p_j = p = 1/C$. Consequently, the size of the clusters is binomially distributed with parameters S and $1/C$, with an average cluster size of S/C.

We follow (5) to calculate the total controller to controller traffic load. Since now we have random topology, we express the mean value, $\bar{T}_{c2c}^{ER} = \sum_{i=1}^{C} \sum_{j=1, j\neq i}^{C} B_i P_{ij} = \bar{B}_{c2c} \bar{P}_{c2c}$.

Again we need the topology specific values of L and L_d to express $\bar{B}_{c2c}$. Since now we have a random topology, the L and L_d values are random numbers. As B_{c2c} depends linearly on L and L_d, we use the average values to express $\bar{B}_{c2c}$. Let us introduce $\bar{k} = (S-1)p$ as the average number of links a switch has. The expected number of intra-cluster links of a switch s is

$$\bar{L}_s = E[L_s] = \sum_{q=1}^{S} \sum_{n=0}^{S-1} E[L|q,n] = \frac{(\frac{S}{C}-1)\bar{k}}{S-1}, \tag{18}$$

where n denotes the number of links the switch has, and q the size of the cluster it belongs to, and both of these random variables have Binomial distribution $B(S-1,p)$ respectively $B(S,1/C)$. $E[L|q,n]$ is the expected number of intra-cluster links under given n and q values, with distribution $B(n,q/S)$.

Then, for the average number of intra-cluster respectively inter-cluster edges in the entire network we get

$$\bar{L} = \frac{S}{2}\bar{L}_s = \frac{S(S_C-1)\bar{k}}{2(S-1)}, \tag{19}$$

$$\bar{L}_d = \frac{S\bar{k}}{2} - \bar{L} = \frac{SS_C(C-1)\bar{k}}{2(S-1)}, \tag{20}$$

where $S_C = S/C$ as before.

Now let us consider the average length of the transmission paths between the controllers. We use the results of [21], that gives the CCDF of the path length between any two nodes in an ER graph as

$$\bar{F}(x) = \exp\Big[\frac{\bar{k}^x}{S-1}\Big], \tag{21}$$

and the average path length, which is also the average c2c path length as

$$\bar{P}_{c2c}^{ER} = \frac{\ln(S)-\gamma}{\ln\bar{k}} + \frac{1}{2}, \tag{22}$$

where $\gamma = 0.5572$ is the Euler constant.

Comparing these parameters to the ones in the regular grid, we see that now a larger ratio of links are inter-cluster links, which slightly increases the generated control traffic according to Table I. Most importantly, however, the average path lengths have very different characteristics for the two topologies. In the grid topology it increases with $\sqrt{S}$, while in the random graph only with $\ln S$, which shows already that the total controller to controller traffic load will be lower under the random graph topology. We see as well that the number of neighbors $\bar{k}$ has stronger effect in the grid topology.

The final expression of $\bar{T}_{c2c}^{ER}$ is

$$\begin{aligned}\bar{T}_{c2c}^{ER} = &\left(\frac{\ln(S)-\gamma}{\ln\bar{k}} + \frac{1}{2}\right)\left(\frac{S}{S_C}\left(b_f^0\left(\frac{S}{S_C}-1\right) + b_s^0\right)\right.\\ &+ S\left(b_{f-}^s\left(\frac{S}{S_C}-1\right)^2 + (b_{s-}^s + b_f^s)\left(\frac{S}{S_C}-1\right) + b_s^s\right)\\ &+ \left(\frac{(S_C-1)\bar{k}}{2}\right)\left(b_{f-}^l\left(\frac{S}{S_C}-1\right)^2 + (b_{s-}^l + b_f^l)\left(\frac{S}{S_C}\right.\right.\\ &\left.\left.\left.-1\right) + b_s^l\right) + \frac{(S-S_C)\bar{k}}{2}\left(2b^d + \left(\frac{S}{S_C}-2\right)b^e\right)\right).\end{aligned} \tag{23}$$

As expected, the increase of the traffic in the number of switches S is lower than in the case of the regular grid, but still the traffic scales with $S^3\ \ln(S)$.

Finally, let us consider the switch to controller control traffic load, which depends on the average switch to controller path length. Let X_i denote the random variable representing the path length from a node to controller i, with CCDF given in (21), and let $Y = \min(X_1, X_2, ...X_c)$ be the path length to the closest controller. Then

$$\bar{F}_Y(x) = (\bar{F}(x))^C = \exp\Big[\frac{C\bar{k}^x}{S-1}\Big], \tag{24}$$

and following [21] we can estimate the mean value of Y, $\bar{P}_{s2c}^{ER}$ as

$$\bar{P}_{s2c}^{ER} = \frac{\ln(\frac{S}{C})-\gamma}{\ln\bar{k}} + \frac{1}{2} = \frac{\ln(S_C)-\gamma}{\ln\bar{k}} + \frac{1}{2}. \tag{25}$$

Since P_{s2c} is proportional to $\ln(S_C)$, the increase in cluster size does not increase the path length significantly. This is a good phenomena as shorter path length guarantees low latency and less traffic in the network. We also see that the positive effect of increasing $\bar{k}$ will be small at large $\bar{k}$ values, and $\bar{k}$ does not change the characteristics of the scaling in network size.

Combining (4) and (25) we can write total switch to controller traffic in the network as

$$\bar{T}_{s2c}^{ER} = \lambda G\left(1-\frac{1}{S_C}\right)\left(\frac{2(\ln(S_C)-\gamma)}{\ln\bar{k}} + 1\right), \tag{26}$$

that is, overall the switch to controller traffic increases with the cluster size, however, this increase is only logarithmic. Similarly to the regular grid, the traffic is proportional to the intensity and the size of the switch to controller queries.

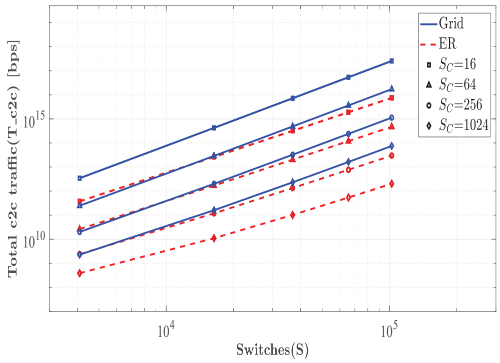

Fig. 5. Controller to controller traffic in regular grid and ER networks as a function of the number of switches S, and for different clusters sizes S_C ($\bar{k} = 4$).

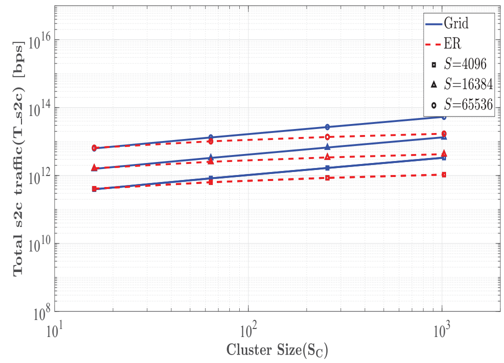

Fig. 7. Switch to controller traffic in regular grid and ER networks as a function of the cluster size S_C and for different number of switches S ($\bar{k} = 4$, $G = 128$Bytes and $\lambda = 25$kpps).

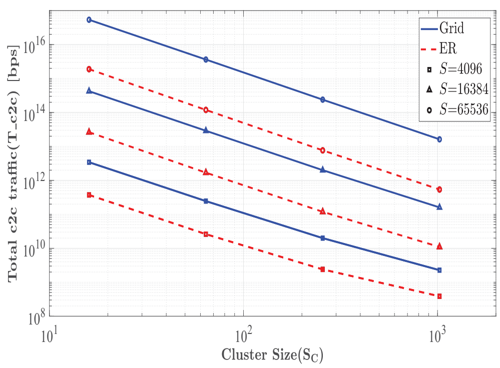

Fig. 6. Controller to controller traffic in regular grid and ER networks, as a function of the cluster size S_C, and for different number of switches S ($\bar{k} = 4$).

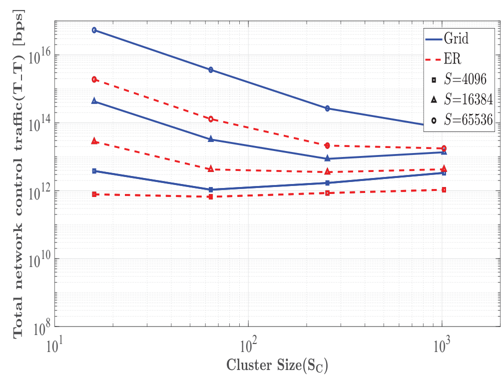

Fig. 8. Total control traffic in regular grid and ER networks as a function of the number of switches S, and for different cluster sizes S_C ($\bar{k} = 4$, $G = 128$Bytes and $\lambda = 25$kpps). Optimal cluster sizes are marked.

V. Numerical examples

In this section we use the expressions derived in Section IV to evaluate the effect of the network size, the cluster size, the intensity of the switch to controller traffic, and the network topology on the value of the control traffic emerging in the network, and discuss the selection of the preferable number of controllers. To be able to evaluate scalability, we consider network sizes in the range of $10^5 - 10^5$ switches, which corresponds to networks of a large number of autonomous systems. Unless otherwise noted, we consider $\bar{k} = 4$, s2c control packet size of $G = 128$Bytes and per switch packet generation rate $\lambda = 25$kpps [14]. The parameter values of the ONOS control traffic are given in Table I.

Figure 5 shows total $c2c$ traffic in the network as a function of the number of switches, for a given cluster sizes for grid and ER networks, on a log-log plot. It reflects well the analytic results on the scaling of the traffic in S. We note that the gradient, giving the exponent of S, is only slightly lower in the ER graph, however, the absolute values differ in ca. one order of magnitude. We see as well that the cluster size has significant effect. The traffic level is very high in general, in the order of terabits per second.

Let us next evaluate the amount of the controller to controller traffic as a function of the cluster size. Figure 6 shows that the decrease of the control traffic is similar in the grid and ER topologies, and is dominated by the parts of (11) and (23) with quadratic decrease, which would motivate the use of large clusters in both topologies.

In contrast, we expect that decreasing cluster size increases the switch to controller traffic. As we see on Figure 7, this increase is significant in the case of the grid topology where the switch to controller path length is proportional to $\sqrt{S_C}$, while less dominant in the ER case, where the path length to the closest controller increases with $\ln S_C$.

Figure 8 gives the total traffic and optimum cluster size for both topologies. As the c2c traffic is dominating in large networks, the optimal cluster size is also large. We see that in middle-sized networks the traffic decreases significantly at smaller than optimal cluster sizes, while there is little penalty if the cluster size is too large. In small networks, under and over dimensioned cluster sizes have similar effect. In general, the cluster size that minimizes the total traffic is smaller

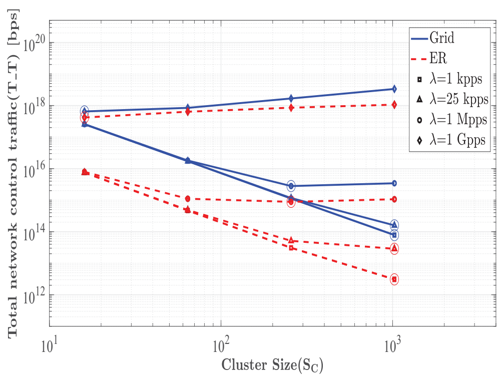

Fig. 9. Total control traffic in regular grid and ER networks as a function of the cluster size S_C, and for different switch to controller traffic intensity values λ (S = 16384, $\bar{k}=4$ and G = 128Bytes). Optimal cluster sizes are marked.

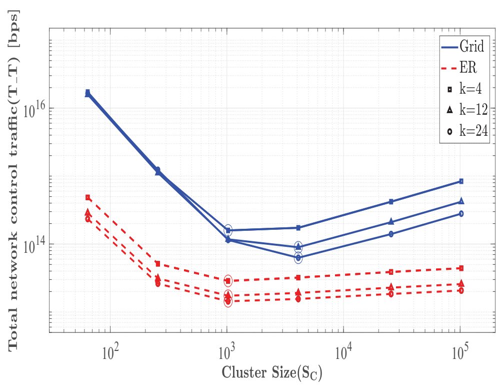

Fig. 10. Total control traffic in regular grid and ER networks as a function of the cluster size S_C, but for different degree values $\bar{k}$ (S = 102400, G = 128Bytes and $\lambda = 25$kpps)). Optimal cluster sizes are marked.

in smaller networks, and, for the considered parameters, the optimal values are similar for the two network topologies.

Figure 9 compares the total traffic for different switch to controller query rates λ and fixed network size $S = 16384$. The figure shows that the intensity of the s2c traffic has significant effect not only on the traffic levels, but also on the preferred cluster size. High λ allows only smaller clusters, while under low traffic intensity larger clusters can decrease the control traffic significantly. Therefore the network should be configured, or even reconfigured according to the s2c traffic.

Finally, on Figure 10 we evaluate the effect of the network density, that is, the average node degree $\bar{k}$, considering k values from 4 up to 24. The analytic results already showed us that $\bar{k}$ does not affect how the traffic volume depends on the network size S and on the cluster size S_C, but scale the volume of the traffic. Increased network density has the negative effect of increasing the amount of information that needs to be exchanged in the control traffic, but it also have the positive effect of shorter transmission paths. In the grid network, these two are in balance when the c2c traffic dominates, and we see the positive effect of increased $\bar{k}$ only when the s2c traffic becomes significant. In the ER network the positive effect is visible for all cluster sizes, but the gain diminishes already at around $\bar{k} = 24$, even for the considered large network.

VI. Conclusions

In this paper we have modelled the volume of the control traffic in distributed SDNs, when ONOS is used for maintaining the shared view of the network. We evaluated the effect of the network size, the number of controllers in the network, and the network topology. We have seen that regular grid networks will have high control traffic due to the long transmission paths, but even random graphs with small-world property suffer from heavy control traffic. We have demonstrated that there is an optimal cluster size that minimizes the volume of the control traffic, and that this cluster size depends significantly on the intensity of the switch to controller traffic, but is often less sensitive to the network topology itself. Similar studies could also help the network design in specific application areas, like data center networking [22]

We have shown that the density of the network have contradicting effects on the control traffic, and little is gained in well-connected networks. Our results consider full mesh connectivity among the controllers. This ensures the fast convergence of the anti-entropy algorithm, but leads to heavy control traffic, due to the long controller to controller paths. A further optimization of the ONOS protocol thus could employ different connection patterns, to trade off consistency and control traffic load.

References

[1] B. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, "A survey of software-defined networking: Past, present, and future of programmable networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1617–1634, March 2014.

[2] M. Caria, A. Jukan, and M. Hoffmann, "A performance study of network migration to sdn-enabled traffic engineering," in *Proc. of IEEE Globecom*, 2013.

[3] V. Yazici, M. O. Sunay, and A. O. Ercan, "Controlling a software-defined network via distributed controllers," in *Proc. of the NEM Summit*, 2012.

[4] A. S. Muqaddas, A. Bianco, P. Giaccone, and G. Maier, "Inter-controller traffic in ONOS clusters for SDN networks," in *Proc. of IEEE ICC*, May 2016.

[5] S. Lange, S. Gebert, T. Zinner, P. Tran-Gia, D. Hock, M. Jarschel, and M. Hoffmann, "Heuristic approaches to the controller placement problem in large scale SDN networks," *IEEE Transactions on Network and Service Management*, vol. 12, no. 1, pp. 4–17, March 2015.

[6] G. Yao, J. Bi, Y. Li, and L. Guo, "On the capacitated controller placement problem in software defined networks," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1339–1342, Aug 2014.

[7] B. P. R. Killi and S. V. Rao, "Capacitated next controller placement in software defined networks," *IEEE Transactions on Network and Service Management*, vol. 14, no. 3, pp. 514–527, Sept 2017.

[8] A. Sallahi and M. St-Hilaire, "Optimal model for the controller placement problem in software defined networks," *IEEE Communications Letters*, vol. 19, no. 1, pp. 30–33, Jan 2015.

[9] N. Perrot and T. Reynaud, "Optimal placement of controllers in a resilient SDN architecture," in *Proc. of IEEE Design of Reliable Communication Networks (DRCN)*, 2016.

[10] F. J. Ros and P. M. Ruiz, "On reliable controller placements in software-defined networks," *Computer Communications*, vol. 77, pp. 41 – 51, March 2016.

[11] T. Zhang, P. Giaccone, A. Bianco, and S. D. Domenico, "The role of the inter-controller consensus in the placement of distributed SDN controllers," *Computer Communications*, vol. 113, pp. 1 – 13, Sept 2017.

[12] P. Megyesi, A. Botta, G. Aceto, A. Pescapè, and S. Molnár, "Available bandwidth measurement in software defined networks," in *Proc. of ACM Symposium on Applied Computing*, 2016.

[13] A. Bianco, P. Giaccone, A. Mahmood, M. Ullio, and V. Vercellone, "Evaluating the sdn control traffic in large isp networks," in *2015 IEEE International Conference on Communications (ICC)*, June 2015.

[14] D. Levin, A. Wundsam, A. Feldmann, S. Seethamaran, M. Kobayashi, and G. Parulkar, "A first look at OpenFlow control plane behavior from a test deployment," Technische Universitt Berlin, Fakultt Elektrotechnik und Informatik, Tech. Rep., 2011.

[15] P. Berde, M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, and W. Snow, "ONOS: towards an open, distributed SDN OS," in *Proc. of ACM Worshop on Hot Topics in Software Defined Networking*. ACM, 2014.

[16] Z. K. Khattak, M. Awais, and A. Iqbal, "Performance evaluation of OpenDaylight SDN controller," in *Proc. of IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, 2014.

[17] A. Panda, C. Scott, A. Ghodsi, T. Koponen, and S. Shenker, "CAP for networks," in *Proc. of ACM SIGCOMM workshop on Hot Topics in Software Defined Networking*, 2013.

[18] D. Ongaro and J. Ousterhout, "In search of an understandable consensus algorithm," in *Proc. of USENIX Annual Technical Conference*, 2014.

[19] R. A. Golding and D. D. Long, "Simulation modeling of weak-consistency protocols," *Network Systems Design*, Jan 1999.

[20] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.

[21] J. A. H. Agata Fronczak, Piotr Fronczak, "Average path length in random networks," *Physical Review E, vol. 70, Issue 5, id. 056110*, pp. 1–7, 2004.

[22] W. Xia, P. Zhao, Y. Wen, and H. Xie, "A survey on data center networking (dcn): Infrastructure and operations," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 640–656, 2017.