

Thesis “Control Plane in Software Defined Networks and Stateful Data Planes”

Candidate: Abubakar Siddique Muqaddas

Tutor: Paolo Giaccone

Answer to the reviewers' comments

First of all I wish to thank the reviewers for their time and effort in helping to improve this thesis. Thanks for all the valuable suggestions and comments, which are considered and addressed in detail. In the following paragraphs the authors present the modifications performed on the manuscript following the reviewers' suggestions.

1 Reviewer: Franco Callegati

Content

Is the title appropriate?: Yes
Is the abstract clear?: Yes
Is the thesis clear on what was done and why?: Yes
Is the order of presentation satisfactory?: Yes

Presentation

Is this dissertation technically sound?: Yes
Does it describe original work?: Yes
Are the references adequate?: Yes
Should some material be omitted?: No
Should some material be added?: No

Overall Rating

Based on your experience as a thesis evaluator, how would you rank this thesis among all theses you have evaluated so far?: B (between 10% and 20%)
Recommendation: Minor Revisions

Comments

Please comment on the originality, methodology, results and perspectives of the work described in this dissertation:

The work presented in this thesis is at the start of the art of the research in the specific domain and, at the best of my knowledge, original. The research work described in the dissertation spans from theoretical to more practical methodologies, which is an added value in this field, in my personal view. The work presented in the second part is more theoretical, with application of optimization modeling. The former part presents a research work which is based on a more practical activity encompassing protocol and traffic analysis, still with some theoretical contributions in the attempt to identify curves that, fit to the experimental results, may provide some engineering guideline for these architectures.

In terms of scientific results I believe the thesis is adequate and will lead to a successful defence.

I just have a few comments suggesting minor modifications.

Comment 1: I suggest a full check of the writing, also by means of automatic tools, to solve some language issues I noticed. For instance, at page 25, subsection 2.4.1 line 10 the text “Tombstone traffic is due the antientropy protocol ...” should be “Tombstone traffic is due TO the antientropy protocol”. At page 51 section 3.1 line 22 the sentence “Although legacy GMPLS/Path Computation Element (PCE) architectures already offered a fully separated control plane from the control plane ...” is not clear.

Reply to Comment 1:

We thank the reviewer for suggesting to fix the language/grammar issues. All these have been fixed in the updated version of the thesis.

Comment 2: Also sometimes in the text there is a reference to “this article” that should be “this thesis”, probably a leftover from previous works, like at page 97 four lines from the end. Please fully check for the article word in the text and modify accordingly.

Reply to Comment 2:

We thank the reviewer for indicating this error. This has been fixed in the updated version of the thesis.

2 Reviewer: Francesco Musumeci

Content

Is the title appropriate?: Yes
Is the abstract clear?: Yes
Is the thesis clear on what was done and why?: Yes
Is the order of presentation satisfactory?: Yes

Presentation

Is this dissertation technically sound?: Yes
Does it describe original work?: Yes
Are the references adequate?: Yes
Should some material be omitted?: No
Should some material be added?: Yes

Overall Rating

Based on your experience as a thesis evaluator, how would you rank this thesis among all theses you have evaluated so far?: B (between 10% and 20%)
Recommendation: Minor Revisions

Comments

Please comment on the originality, methodology, results and perspectives of the work described in this dissertation:

The thesis provides various contributions concerning the interfaces of a SDN control plane and as well addresses the optimal placement of state variables in stateful data planes. I particularly appreciated the way the chapters have been organized and the overall context of the thesis, as the various contributions properly fit the entire set of the interfaces in a SDN controller. This has helped to make the thesis a complete and original contribution. The thesis includes a wide experimental activity, especially Part I, whereas in Part II the development of a mathematical ILP-based model provides stronger methodological validity to the work. Despite it is technically sound, the mathematical modeling provided in Part I appears much "simpler" than that in Part II. Overall, the adopted methodology is fair. Results are presented in a satisfactory way, although some revision should be brought (see below).

The most important changes and suggestions to improve thesis quality are reported below.

2.1 Conceptual Issues

Comment 1: In Chapter 2, it is often stated that the model provided allows to extend/-generalize the considerations to *any* topology; this statement is in my opinion too strong as it is performed by only considering results for very simple topologies (isolated/linear/star); at least the author should provide an analysis with more complex (i.e., meshed topologies) and under different scenarios (i.e., nodes under the controllers domains).

Reply to Comment 1:

We thank the reviewer for pointing out the claim that the presented model in Chapter 2 must be supplemented with analysis of more complex topologies. In the light of this, the claim has been revised that the presented model is expected to be applicable to a broad set of topologies. We have corrected this claim in Sec. 2.1.2 and Sec. 2.4.

Comment 2: Eq (3.3) seems not correct to me. I'd express it as $\max(4c_i + RTT_i/2)$. As an example, consider the case when $c_1=c_2=10$, $c_3=0$, and $RTT_1=RTT_2=1$, $RTT_3=2000$; with the provided formula the value of t_{N-NC} would be 1030, whereas with "my" formula it would be 1000. In other words, I think the term to compare is the *joint* contribution of RTT and c_i , and the two terms should not be considered in isolation when extracting the $\max()$.

Reply to Comment 2:

We thank the reviewer for suggestion of jointly considering the contribution of RTT and c_i . However in the case of N-NC, we are assuming that all operations are not performed at each node back to back; rather, the mentioned 4 operations are performed sequentially. Each of these 4 operations (tear down L_3 , tear down L_1 , setup L_1 and setup L_3) is first completed before moving on to the next operation at all nodes. Hence, for the last operation, the term $\max_i(c_i + RTT_i/2)$ is added separately.

Comment 3: eq (3.4) implicitly assumes that the value for c_i considered in the OF/IA-NC case is the same than the other two cases; in my opinion this is not 100% fair, as performing simultaneous operations requires, e.g., that the same processing resources are shared, thus leading to a higher configuration time per single operation.

Reply to Comment 3:

We thank the reviewer for this comment and acknowledge that since multiple wavelengths are being configured simultaneously, this may lead to different configuration times for ASY vs OF/IA-NC as per Fig. 3.4(b). Taking this into consideration, at the end of Sec. 3.4, we present another result that takes into account different configuration time for ASY/N-NC and OF/IA-NC case. We present an example of the scenario in Fig 3.1 and use the configuration times of Fig. 3.4(b). Even though the reduction is less as compared to the claimed 75% in an ideal case, still due to the bundling of operations in OF/IA-NC, a considerable reduction in lightpath disruption time is still visible.

Comment 4: In Section 3.5.3 I would have expected results also (actually, mainly) in terms of disruption time, so as to validate the analytical model provided in section 3.4

Reply to Comment 4:

We thank the reviewer for his observation of the unavailability of results of the disruption time. The explanation for this is now provided in the last paragraph of Sec. 3.5.2.

Comment 5: o.f (5.1): the second term of the o.f. is an "indirect" way of minimizing the nr of states copies. Is this exactly equivalent to minimizing directly variable $\hat{C}_s$. The author should provide a quantitative comparison for this, or at least motivate why the two objectives are equivalent. The same applies to (5.2).

Reply to Comment 5:

We thank the reviewer for indicating this ambiguity. In fact, the reviewer is correct in saying that the second term of the objective function in (5.1) is just to minimize the state synchronization traffic. This may indirectly minimize the number of copies but does not necessarily imply it. The motivation behind adding the synchronization traffic to the objective function was that it is the same type of quantity as the data traffic; whereas adding the number of copies $\hat{C}_s$ is not homogeneous. Hence, we revise the claim in Chapter 5 that we are minimizing the number of copies, and replace it with "minimizing" the state synchronization traffic in Secs. 5.1 and 5.3.

Comment 6: The complexity of the ILP formulation should be discussed; how does it increase with problem instance?

Reply to Comment 6:

We thank the reviewer for suggesting to discuss the computational complexity of ILP model. The updated version of thesis now includes this in Sec. 5.3.2 and Appendix A.

2.2 Formalisms throughout the paper

Comment 1: Page (P) 21: it is stated that "In each experiment, we start the controllers and then we wait until the initial transient phase ends"; the "initial transient phase" should be defined more formally, i.e., does it end after certain fixed amount of time? or When the observed inter-controller traffic is within a certain predetermined range? or what else...

Reply to Comment 1:

We thank the reviewer for the comment about the missing formal definition of the "initial transient phase". This has been corrected and a definition of the initial transient phase is provided in Sec. 2.3 in the updated version of thesis

Comment 2: Table 2.1: parameters L and S should be included also in this table (together with all problem parameters)

Reply to Comment 2:

We thank the reviewer for suggesting to include the parameters L and S in the table 2.1. However, the rationale for keeping it separate was to present the notation for 2 controller and 3 controller scenario separately as it is easier to explain and understand while deriving the throughput equations.

Comment 3: P28: B^I is not defined (it's only defined for specific the $x \rightarrow y$ and $y \rightarrow x$ directions)

Reply to Comment 3:

We thank the reviewer for pointing out the missing definition of B^I . This has been defined along with B^L before the description of eq (2.3).

Comment 4: P28: the description before eq (2.6) can be better expressed with a closed-form equation, so as to show how eq (2.7) in Property 1 is obtained.

Reply to Comment 4:

We thank the reviewer for suggesting the inclusion of an equation before eq (2.6) to show how eq (2.7) is obtained. In the updated version of the thesis, the "new" eq (2.6) along with a description is used for this purpose.

Comment 5: P29: it is not clear what is the relation between weight of L_{xy} in eq 2.7 (0.63) and the one for L_{AB} in eq of B_{TOT} (-0.48)

Reply to Comment 5:

We thank the reviewer for indicating the lack of clarity. To rectify this, the derivation of B_{TOT} is presented in eq 2.9 which explains the link between L_{xy} and L_{AB} .

Comment 6: P57: "However, this worst-case configuration time can be estimated with some error...". Can the author provide a quantitative (theoretical) upper bound?

Reply to Comment 6:

We thank the reviewer for his observation regarding the worst-case configuration time of an optical time. However, in the thesis, this is stated that it is highly dependent on many factors which may not be limited to the current load of the agent at the optical node, aging issues, different firmware/software implementations, control plane protocols as well as random behaviors. To further elaborate, we have mentioned the reference to this statement along with the reference to Fig. 3.4(b) which is an example of the configuration time for a WSS.

Comment 7: Table 3.1: I do not understand how the numbers in column "No. Channels" have been calculated. Considering that it is stated that the channels have 50GHz spacing and considering the 1st and last channels in the table I'd calculate this number as $(194-192.8)\text{THz}/50\text{GHz}=24$. What am I missing? Please clarify.

Reply to Comment 7:

We thank the author for pointing the missing link between number of channels as a function of the channel spacing. In the updated version of thesis, we have included the explanation for the number of channels in Sec. 3.5.2. This is accompanied by an illustration in Fig. 3.8 to further clarify.

Comment 8: Table 5.1: all the inputs are not actually "variables" but rather parameter-s/sets; moreover, include ALL the inputs in the summary table (e.g., I don't see V , N , E , c_e in the table)

Reply to Comment 8:

We thank the reviewer for indicating the correct term for Table 5.1 and the missing inputs. In the updated Table 5.1, the caption has been changed and the missing inputs have been added.

Comment 9: Table 5.1 vs 5.2: known (i.e., given inputs) and variables should be ALL (e.g., Y_{smne} is not there) included in two different well-distinguishable tables. Use of term "variables" also for the first table is confusing.

Reply to Comment 9:

We thank the reviewer for pointing out the missing variable Y_{smne} . In the updated Table 5.2, this has been fixed.

Comment 10: P109: the "sufficiently large" M is not rigorous definition; a lower bound for M should be mathematically provided, i.e., according to the sets/variables it is supposed to bound

Reply to Comment 10:

We thank the reviewer for suggesting to include the definition of M . In the updated version of thesis, equations (5.17) and (5.18) have been provided to mathematically define M .

Comment 11: P110: "most central" is way too generic; the author refers to betweenness centrality later; this should always be the rigorous term to use.

Reply to Comment 11:

We thank the reviewer for indicating this. It has been corrected in the updated version of thesis, and the term "most central" has been replaced with "node with the highest betweenness centrality" in Secs. 5.2.3 and 5.4.

Comment 12: Sec. 5.4.1: the author should provide an illustrative example of the unwrapped Manhattan and Watts-Strogatz.

Reply to Comment 12:

We thank the reviewer for pointing out the missing illustrations. In the updated version of thesis, the two topologies are now illustrated and depicted in Sec. 5.4.1.

Comment 13: Sec 5.4.1: it is mentioned that results are obtained with "very small" 95% confidence intervals; this seems too vague; can the author provide a number for that? e.g., in fig 5.2 (in the nr of copies result, left-most point) there is a visible confidence interval, differently from the other points. A similar comment applies to section 5.5.2.

Reply to Comment 13:

We thank the reviewer for indicating the vagueness of confidence intervals. The maximum accuracy for the results is now provided in Sec 5.4.1 (in all cases within 4.2% accuracy). The author would like to inform that in some cases, the confidence intervals are so close to the mean value that they are unable to be seen from the naked eye in the graph, however they are still within 4.2% of the mean value.

Comment 14: P118: to my understanding, the value 0.5221 comes automatically once the topology is given (it is stated that "Similarly for the synchronization traffic, the average distance is computed. However, no random points are generated since the copies have a fixed position in the unit square"); if so, it should be expressed in a closed-form formula as a function of the number of copies (and maybe other topological parameters).

Reply to Comment 14:

We thank the reviewer for pointing the closed form formula for the $\hat{d}_{sync}$. In the updated version of thesis, a closed-form formula has been provided in equation (5.28), which is used to compute the average synchronization distance shown in Fig. 5.10.

2.3 On the results and graphics

Comment 1: Fig 2.7: it is not clear if the transient phases depicted are affected by the available capacity between the controllers; can the author provide a comment on this (possibly, in a quantitative way)?

Reply to Comment 1:

We thank the reviewer to inquire about the impact of available capacity between the controllers on the transient phases of the traffic. To clearly capture the transient phases and the steady state bandwidth, we assume an inter-controller traffic channel to have a huge bandwidth. Here, since experiments are carried out using LXC, all traffic between ONOS controllers is passing through linux bridges and is processed by the linux kernel. Consequently, the available capacity between the controllers is limited on the resources of the server hosting the LXC containers. Here, we are not imposing any limits on the bandwidth of the linux bridge; consequently justifying the claim for the huge bandwidth.

Comment 2: Figs 2.8 and 2.9: do these results represent the steady-state values after the transient phase occurring when inserting nodes in the topology? In other words, are they comparable with 88 kbps in Fig 2.7? If yes, I do not understand why there is a difference between this value (88) and the one in Fig 2.9(a) for $S = 10$, which seems to be above 100 kbps and (to my understanding) represents the same case of that in Fig 2.7.

Reply to Comment 2:

The reviewer has correctly pointed out the mistake. In fact, the author acknowledges the error and has corrected the description for Fig. 2.7 in Sec. 2.4.1. The case in Fig. 2.7 in fact refers to the case for $S = 4$ and not $S = 10$.

Comment 3: Figs 2.8 vs 2.9: it's worth noting (and explaining why) the presence of links have a small effect in the direction $A \rightarrow B$ and no effect at all in the direction $B \rightarrow A$; can the author motivate this?

Reply to Comment 3:

The reviewer has pointed out that in the throughput curves of Figs. 2.8 and 2.9, the effect of links on the throughput is not visible for the case $B \rightarrow A$. This effect is not noticeable since the effect is very small, as pointed out in eq (2.4), where $b_{B \rightarrow A}^l = 0.11$ kbps. The effect is very small, since switches are only added to controller A, and the inter-controller traffic for $A \rightarrow B$ not just includes the contribution of the anti-entropy protocol, but also the update on the topology store of controller A due to topology discovery via LLDP. However, for B to A , the inter-controller traffic for the topology store only includes the contribution of the anti-entropy protocol, thus the value of $b_{B \rightarrow A}^l$ is small and unnoticeable in the throughput curves.

Comment 4: Fig 2.18-left: there seems to be a slight decrease from the first to the second set of points? Can this be motivated?

Reply to Comment 4:

We thank the reviewer for pointing out the deviation in the set of points in the Figure. The decrease is due to an initial error during measurement, which is averaged when a huge number of flows is installed.

Comment 5: Fig 3.3 shows that there is a difference in the configuration times in sub-figures (b) and (c), but this is not the case for subfigure (a). Can the author motivate this? This might affect results.

Reply to Comment 5:

We acknowledge the reviewer's comment that the configuration times between ASY and N-NC cannot be observed in Fig. 3.5. This is due to the fact that N-NC, as shown in (3.3), includes the term $RTT/2$; whereas ASY, as per (3.1) and (3.2), includes the round trip time as $3.5 \times RTT$. Thus the difference of RTT between ASY and N-NC is $3 \times RTT$, which amounts to 30ms in Fig. 3.5 and is thus negligible as compared to the overall disruption time. Hence, there is a difference in the configuration times. This has been updated in the Sec 3.4 of thesis.

Comment 6: Fig 3.5: the claimed 75% reduction is not constant, is it the maximum? please clarify.

Reply to Comment 6:

We thank the reviewer for indicating the claimed lightpath swapping reduction is not constant. In the updated version of thesis, wherever we have claimed 75% reduction, we have claimed it is the minimum. It is also evident from Figs. 3.5 and 3.6.

Comment 7: Fig 5.2: it would be nice to see for which value of λ_s the number of copies goes to 1: this represents a "tolerance" threshold for synchronization traffic.

Reply to Comment 7:

We thank the reviewer for this interesting suggestion. The updated version of the thesis now includes a result in Fig. 5.4 showing the value of λ_s for which the number of copies goes to 1, along with an explanation of the result in Sec. 5.4.1.

2.4 Minor changes

2.4.1 On the writing

Comment 1:

P4, last line: "...enhanced mobile broadband (eMBB) with data rates up to 10 Gbps..."
→ "...enhanced mobile broadband (eMBB) with data rates up to 10 Gbps to end users..."
P5, sect. 1.2: "The focus ...are"→"The focus...is"
P5, sect. 1.2: "It refers to the each..."→"It refers to each..."
P8, 1st paragraph: statement "The key idea to allow operators to... interconnection." has no main verb. Maybe substitute "based" with "is based"
P9, sect 1.2.4, 4th line: "for" seems unnecessary

Reply to Comment 1:

The author thanks the reviewer for indicating the errors. These have been adjusted appropriately as suggested.

Comment 2: Throughout Chapter 2 (but this applies in general): the term "bandwidth" is used several times (e.g., in Fig 2.3 and in the text) referring to the intercontroller traffic; I'd suggest to use "traffic" or "intercontroller traffic" instead (to my view, "bandwidth" usually refers to the available capacity).

Reply to Comment 2:

We thank the reviewer for indicating the ambiguity related to the term "bandwidth". The reviewer is correct when referring to the actual definition of bandwidth: the available capacity. In light of this, throughout Chapter 2, the term bandwidth has been appropriately replaced with the term throughput. The term "bandwidth" is retained in Chapter 2 wherever it actually refers to the available capacity.

Comment 3: LLDP acronym is not defined in the text and not even included in the list of acronyms.

Reply to Comment 3:

We thank the reviewer for noticing the missing definition and acronym of the Link Layer Discovery Protocol (LLDP). In the updated version of the thesis, the acronyms have been updated and the definition of LLDP is provided in Sec. 2.4.

Comment 4:

P25: "Tombstone traffic is due the" -> "Tombstone traffic is due to the"

P38, sect. 2.5.2, 2nd line: repetition of "possible"

P46: Sentences should not start with a "direct" reference, i.e., use "Ref. [12] discusses..." instead of "[12] discusses...". This comment might apply to other parts of the thesis

P51, 2nd paragraph, 4th line: "offered a fully separated control plane from the control plane"->"offered a fully separated control plane from the data plane"

P53, section 3.1.2, 3rd line: "Sections. 3.4 and 3.5"->"Sections 3.4 and 3.5" (unnecessary dot)

Reply to Comment 4:

The author thanks the reviewer for indicating the errors. These have been adjusted appropriately as suggested.

Comment 5:

P54: "Note that this operation of lightpath swapping implies a non-negligible disruption time for L3"->"Note that this operation of lightpath swapping implies a non-negligible disruption time for L3 and L1". In fact fig 3.3a shows the disruption also for L1, then doing make-before-break is an exception, as stated later by the author. P56: "...does not experience any disruption time."->"...experiences a negligible disruption time." this seems more rigorous to me.

P57: "see Fig. 4(b)"->"see Fig. 3.4(b)"

P57: "Indeed, given the..."->"Indeed, the..."

Reply to Comment 5:

The author thanks the reviewer for indicating the errors. These have been adjusted appropriately as suggested.

Comment 6: P57: please provide a reference for statement "makes its contribution to the TSO inaccuracy negligible compared to optical configuration times which are in the order of seconds"

Reply to Comment 6:

We acknowledge the reviewer's comment and the sentence is now referenced.

Comment 7:

P62, last line: "though"→"through"

P63: "stablish"→"establish" (this appears several times)

P65: "quantized by"→"confirmed in"

P65: "one order of...(in dB) lower" → "10 dB less"

P65: "number of optical channels increase" → "increases"

P75, 2nd line: unnecessary ")" after "VNFs"

P84: "of a network service consisting of one VNF." → "of a network service consisting of one VNF per island."

P108: "...and f exploit..."→"and f exploits..."

P119: "the number of copies decrease"→"the number of copies decreases"

Reply to Comment 7:

The author thanks the reviewer for indicating the errors. These have been adjusted appropriately as suggested.

2.4.2 On figures/tables clarity and explanation:

Comment 1: Fig 2.3: the title of Y axis is "bandwidth" but it can be more explicitly termed as "intercontroller traffic".

Reply to Comment 1:

We thank the reviewer for indicating the ambiguity related to the term "bandwidth". As stated in a previous comment, throughout Chapter 2, the term bandwidth has been appropriately replaced with the term throughput.

Comment 2: Fig 2.8-2.9 (this comment might apply also to other figures): captions should be made more significant as they do not describe what is in the charts (i.e., the intercontroller traffic) but rather they are describing some details of the scenario for which the results are obtained (i.e., the topology, the fact that A is master)

Reply to Comment 2:

We thank the reviewer for suggesting to make the captions more significant by adding in more details. In the updated version of thesis, the captions for Figs. 2.8, 2.9, 2.12, 2.13, 2.15, 2.16, 2.19 and 2.21 include more details.

Comment 3: P35: is there any particular reason for considering the 261th ISP as an example? If so, state it, otherwise say explicitly that it is just an example.

Reply to Comment 3:

We thank the reviewer for this insight. In the updated version of thesis, the statement has been modified as per the reviewer's suggestion.

Comment 4: P53, introducing Fig 3.1: more introductory details should be provided, e.g., in the form "in the figure we show a sample topology with 5 nodes and 5 links representing (...), where 4 lightpaths are routed and require X, Y, Z, W FS, respectively..."

Reply to Comment 4:

We thank the reviewer for indicating that more details should accompany Fig. 3.1. In the updated version of thesis, this is included in Sec. 3.2.

Comment 5:

Fig 4.2: a legend for the yellow/purple/red times can be useful

Fig 4.2: in some arrows it should be specified which operations are actually requests, e.g., "service composition", "service instantiation" and "service deployment" should be "service composition request", "service instantiation request" and "service deployment request", respectively.

P82: procedures detailed in the text (i.e., "When the 5GUKEx receives the responses from all the relevant islands, it informs the user that the islands are ready to deploy the inter-island NS end-to-end" and "After the inter-island service interconnection is provisioned, the deployment procedure is finished and the end-user is notified") are not reported in the figure 4.2. Please fix.

Fig 4.4: lines (i.e., dashed/dotted...) in the legend do not match with the lines in the chart; please fix.

Reply to Comment 5:

The author thanks the reviewer for indicating the errors. These have been adjusted appropriately as suggested.

Comment 6: Algorithm1: please identify phases 1 to 3 mentioned on page 110 also in the pseudocode.

Reply to Comment 6:

In the Algorithm 1, the phases have been identified on the right side. To improve the readability, all the "phases" are now written in a bold style in the updated version of thesis.