Consolidated Review of

Demystifying the Dark Side of the Middle: A Field Study of Middlebox Failures in Datacenters

1. Strengths:

This paper's findings on middlebox failures are very helpful in the light of the recent attention to middleboxes, their role in the overall network architecture, and various proposals for managing and orchestrating them. Useful take on the kinds of failures that permeate middleboxes.

Study has implications for middlebox design, management techniques and overall data center and application operations. The observation that crash-stop failures aren't as common as grey-failures has ramifications for designing robust network protocols.

The large-scale dataset used in this paper is quite unique and interesting.

Well-written, thorough paper

2. Weaknesses

Not clear how general the results and takeaways are. The discussion in this paper is highly specific to the topology and sometimes is not convincing. Some of the conclusions being drawn through the paper are not surprising, not new, or are limited in utility (see detailed comments).

Unclear how the failure determination steps are applied. Implications discussed don't seem to tie into the empirical observations.

Some more details about the dataset may be necessary

Section 3.4 is limited in scope.

The root cause analysis section does not provide as strong an analysis as one would expect.

3. Comments

This is a great IMC paper: it tackles an important problem (middleboxes in the cloud); it gathers a novel, comprehensive data; it presents a broad set of measurement insights while also providing a longitudinal view. The writing is crisp and there are no real issues. This is the kind of paper that would be cited many times over in years to come.

Overall, this reviewer likes the topic and the approach that has been chosen by the paper. The overall flow of the paper is nice. The story is well motivated, and the findings are interesting.

The writing can be improved by removing some redundancy though. For example, the findings of the paper are repeated (almost verbatim) four times just in the first two pages of the paper. I wouldn't mind repeating as long as there are new points to be made, or a different perspective is provided, but this is not the case. Or, Section 2.1 defines what firewalls, IDS devices, and load balancers do, which is (IMO) hardly needed for this community.

While I don't have any major concerns, I do have a couple of comments that can help improve the paper:

1. My main concern reading the paper was about the generality of results. Many of your observations seem to be vendor/product line specific. As such it is hard to tell how general your observations and to what extent they apply to other data center settings where middleboxes are applied. That said, this paper's view into a specific middle box deployment is still valuable as this data is simply not available today.
2. I liked section 2 in that it is quite thorough. I really liked the picture showing how the different filters were applied in determining failures. However, the text was quite dense and could use some more structure and less cramming. This section took me forever to read!
3. In the bottom part of Figure 1 the methodology: are these classes mutually exclusive? How the categorization has been conducted? A brief description would be helpful.
4. In Section 2.3, the question of what a failure is interesting and difficult to answer as pointed out by the paper. I understand the challenges involved here, but relying on traffic and operator tickets might significantly underestimate failure events. The paper needs to be more specific about the methodology and its implications on the results presented here. Similarly, the challenges described in Section 2.3 focus on eliminating redundant failures and ignore coverage. Without an explanation of coverage, all the results presented are a lower bound on failures. This is fine for some conclusions reached by the paper, but a clarification and justification is required IMO. The root cause analysis section is a critical part of this paper, as any lessons learned for the purpose of future use would depend on this part. Having said that, I think the paper takes a rather simplistic approach here that might lead to erroneous conclusions. The results presented in Section 6 are interesting, and might have valuable implications for future system design.
5. In table 1, row 2, column 1: you argue that "Making Middleboxes Someone Else's Problem" paper introduces "hardware problems" as the majority of failures, but this is not a valid argument against that paper. Actually, the estimation there shows that most of network administrators believe "misconfiguration" is the root cause not "hardware problems. However, you show connectivity error is the major problem. For obvious reasons (as you confirmed in the last paragraph of page 8, there might be 70% intersection) most of connectivity errors can be as a result of misconfigurations. Thus, my question is that why do you think that your observation is different than others? (3) What does IDPS stand for? (4) The term COV is defined on the page 10, but used multiple times before that.
6. Some discussions don't have enough depth. Thus, the detailed analyses don't propose any general patterns. For instance, not sure what is the benefit of knowing that your specific vendor has updated its software to be more resilient to the failure (section 3.1)? In addition, sometimes even these trivial justifications are left out. For example, in table 2, why does the AFR rate of VPN increases? In summary, since there is no correlation among random variables you are discussing inter-type and intra-type analysis are not useful.

7. The authors show in section 3.1 that middleboxes become more reliable and experience fewer failures in /general/ as they are upgraded to newer versions and older ones reach the end of their product cycles. This is not surprising and this reviewer fails to see how this information (or information about the exceptions to the rule) can be taken advantage of. - The observations about, 1) the failure rates of load balancers and 2) the diminishing returns of network redundancy are not new and have been drawn in [1] as well (including the numbers for the load-balancer observations). [1] Phillipa Gill, Navendu Jain, and Nachiappan Nagappan. 2011. Understanding network failures in data centers: measurement, analysis, and implications. Overall the paper would benefit from a language check and proper sizing of all Figures, etc. Please ensure readability of all aspects of the paper.
8. Section 3.4 is weak and limited in scope, in that the authors have attempted to fit a model to their data, which represents the observations made at only a single provider, and then conclude that the model can approximate real-world middlebox failure data at data-centers.
9. It would be nice to know more about the data set and the environment it was collected for. What applications is the data center running? How much traffic does it see? How many applications having middleboxes interposed in their end-to-end communication flow? What are some of the details of the devices employed, e.g., are they all from the same vendor? Is there vendor heterogeneity in a given device type? How much traffic do different middleboxes/middlebox types process? How many different middleboxes does an average end-to-end traffic flow interact with? It seems to me that knowing these details is useful to understand the overall importance of the observations you make.
10. I found the discussion section to be really bizarre. I don't think the paper would suffer if you took this out entirely! In many cases, there was no clear link between what you are discussing and the empirical results. This is perhaps best illustrated by your discussion of software-defined middlebox networking which has absolutely no link to your measurements. I would have much rather had a section that elaborated on the implications you discuss in Table 1 of your paper.

4. Summary from PC Discussion

Strengths:

- Important topic
- First view of middle box failures in data centers.
- Useful takeaways for practitioners.

Weaknesses:

- Root cause analysis is weak
- Discussion does not tie into measurement results
- Some writing issues; extra analysis needed

On the whole, this is a great paper. No need for much discussion. The weaknesses are not fatal and could be addressed through shepherding.

5. Authors' Response

The main clarifications asked by the reviewers are:

Failure analysis and takeaways:

- **Generality of results:** As with any real-world empirical study, our results are limited by our measurement data. However, we believe it is representative of middlebox deployment in other data centers based on the scale (2k+ devices), diversity (multiple vendors and platforms), duration (two years), and validation from operators specializing in middleboxes.
- **Utility of results:** Table 1 summarizes the key findings and implications of our study: (a) Majority of the failures are grey, (b) hardware errors, misconfigurations, and overloads occur less than expected, (c) there is a broad range of misconfigurations, (d) middlebox redundancy is 67% effective for load balancers and firewalls, and (e) the prevalence of 'few bad apples' effect in a family of middleboxes.
- **Scope of Failure modeling:** Section 3.4 aims to model TTF and TTR for load balancer failures in our study to help guide researchers in improving their design and reliability. We aim to extend our failure modeling to other types of middleboxes in the future, and have clarified it in the text.
- **Root cause analysis:** Section 4 analyzes middlebox failures by using NetSieve (reference [25] in the paper) to infer their root causes from tickets written by operators having expertise in troubleshooting them. Thus, the conclusions we draw are based on aggregating the main operator findings across devices.
- **Implications tied with results:** We have clarified Section 7 by (a) summarizing the key results and adding references to their sections for each discussion point, and (b) tying the implications to Table 1.
- **Intra-device analysis:** In Section 3.1, the observation that a vendor has updated its software to improve failure resilience of a platform can be used to identify outlier devices that still exhibit high failure rate or high downtime despite the update. We have clarified the text.

Details on the dataset and writing:

- **Details on Figure 1 categorization:** Operators tagged each incident to a problem category based on the primary issue experienced by impacted customers.
- **Details on methodology and datasets**: Our measurement data is collected from 2k+ middleboxes across vendors and platforms that carry traffic for a wide range of applications such as web services, search, email, cloud computing, video streaming, and high business impact applications. We ensure 100% coverage of failure events by including all events recorded in network tickets deemed 'actionable' by operators.
- **Cutting repeated text:** We have cut the repetition of our results in Section 1.1. We have kept the middlebox overview text in Section 2.1 to provide completeness for readability.
- **Missing definitions:** We have defined IDPS and COV in Section 1 and 3.1, respectively.
- **Readability of the figures:** We have ensured proper sizing of all the figures.