

Review #: 1

Abstract and Introduction are adequate?	Yes
Conclusions/Future Work are convincing?	Yes
Figures are adequate ? in number and quality	Yes
Improve critical discussion ? validation	Yes
Improve English?	No
Needs comparative evaluation?	Yes
Needs more experimental results?	No
Originality Newness of the ideas expressed	4
Overall Rating Weighted value of above items	4
Paper formatting needs adjustment?	No
Presentation Structure/Length/English	4
References are up-to-date and appropriate?	Yes
Relevance Paper fits one or more of the topic areas?	4
Significance Is the problem worth the given attention?	3
Technical Quality Theoretical soundness/methodology	4
Scale: 1:Lowest Value;6:Highest Value	

Observations for Author

Comparative evaluation is needed. Compare with similar methods or compare the results of system with your tool or without our tool.

Review #: 2

Abstract and Introduction are adequate?	No
Conclusions/Future Work are convincing?	Yes
Figures are adequate ? in number and quality	Yes
Improve critical discussion ? validation	Yes
Improve English?	No
Needs comparative evaluation?	No
Needs more experimental results?	No
Originality Newness of the ideas expressed	2
Overall Rating Weighted value of above items	4
Paper formatting needs adjustment?	No
Presentation Structure/Length/English	4
References are up-to-date and appropriate?	Yes
Relevance Paper fits one or more of the topic areas?	5
Significance Is the problem worth the given attention?	2
Technical Quality Theoretical soundness/methodology	3
Scale: 1:Lowest Value;6:Highest Value	

Observations for Author

In this paper an load generator is described that foremost targets protocols that are in widespread use in networks utilizing technologies implemented in Microsoft's Windows operating system. The targeted applications include Active Directory (a directory service used, e.g., for authentication), Web servers, Remote Desktop Protocol servers (graphical remote access) and file servers utilizing the SMB network file system (used, e.g., for network shares). By utilizing the capabilities of Microsoft PowerShell (a command-line and scripting environment for Microsoft Windows), the presented load generator can mediate complete

authentication procedures, extending the applicability of the approach to services that require proper authentication (i.e., situations where a simple replay of previously captured traffic will be rejected). The presented solution is described to be extensible (which is credible considering the scripting environment utilized) and a small experiment, in which the tool is applied to a virtualized environment comprised of several servers and clients, is conducted to demonstrate the tool's capabilities. It is shown that even a small number of clients can generate considerable server load. The authors concede that the generated load for web servers is highly regular for the time being. In practice, the load inflicted upon web servers is highly dependent on the selection and distribution of the visited dynamically generated pages. Similar considerations may also apply to other loads regarding, e.g., the load of networked storage systems, where access patterns for read and write operations could make a discernible difference. Users that consider using the presented tool may be interested to know how the generated load patterns can be customized to model their real-life scenarios. One motivator for the presented work is described to be failure in cloud infrastructure. However, the services targeted by the tool are most commonly used in private networks (with the important exception of web servers, of course). The experimental setup also mostly resembles a private network scenario. Given the prominent place in the paper (first sentence of the abstract) it is reasonable to assume that some readers may appreciate a short discussion on how the findings can be transferred to cloud infrastructures, which more often than not are of highly distributed nature. Similarly, it should be discussed in what concrete way the presented tool incorporates, implements or enables machine learning and failure prediction. The authors seem to be rather concerned regarding whether the clients employed to generate load remain usable during the production of load. Conventional wisdom suggests that, whenever possible, tests should be conducted outside of the production environment. A few words outlining a scenario where the usable-during-test property is desirable may clear up the confusion some readers may encounter. In summary, the paper is borderline. However, if the paper is accepted the authored should be urged to reduce their claims to the points they really address in the paper!

Review #: 3

Abstract and Introduction are adequate?	Too short
Conclusions/Future Work are convincing?	Yes
Figures are adequate ? in number and quality	Yes
Improve critical discussion ? validation	Yes
Improve English?	Yes
Needs comparative evaluation?	Yes
Needs more experimental results?	No
Originality Newness of the ideas expressed	4
Overall Rating Weighted value of above items	3
Paper formatting needs adjustment?	No
Presentation Structure/Length/English	3
References are up-to-date and appropriate?	Yes
Relevance Paper fits one or more of the topic areas?	5
Significance Is the problem worth the given attention?	4
Technical Quality Theoretical soundness/methodology	3
Scale: 1:Lowest Value;6:Highest Value	

#### Observations for Author

This paper describes a tool to generate network traffic, the so called Distributed PowerShell Load Generator (D-PLG). After a short introduction the author(s) give an overview of the related work in this scope. Thereafter, general aspects of D-PLG are presented and some details of the methodology. The paper concludes by a future work and a conclusions section. The topic of this paper is appropriate to be presented, but the paper itself is in stage of "work in progress" and must be improved prior to publication (weak reject). At the beginning it is not easy to read due to an imprecise problem definition and a short introduction. Although, the authors write about other tools in this scope, a more critical discussion and comparison is desirable. Furthermore, some statements are confusing and have a touch of a subjective perception. For example, "... we believe that these results show that ...", "... our tool generates very realistic web traffic ...". Questions are remaining: How it is measured? What is a very realistic web traffic? The reference to Microsoft community should be quantified by values. To assess the tool performance the paper references to the amount of generated load (e.g., "... fifteen thousand authentication requests over a five minute time period ..."). An interesting question in this context is also, how to handle bursts in one time period (distribution of load). References should not be placed in the abstract, as it can be presented without the paper itself. Finally, for a better readability English must be improved. Many terms are repeating very often: e.g., unfortunately and fortunately (see abstract), "... we demonstrate the validity of our tool and demonstrate ...".

#### Review #: 4

Criterion	Description	Value
Abstract and Introduction are adequate?		Yes
Conclusions/Future Work are convincing?		Yes
Figures are adequate ?	in number and quality	Yes
Improve critical discussion ?	validation	Yes
Improve English?		Yes
Needs comparative evaluation?		Yes
Needs more experimental results?		No
Originality	Newness of the ideas expressed	4
Overall Rating	Weighted value of above items	4
Paper formatting needs adjustment?		Yes
Presentation	Structure/Length/English	4
References are up-to-date and appropriate?		Yes
Relevance	Paper fits one or more of the topic areas?	3
Significance	Is the problem worth the given attention?	4
Technical Quality	Theoretical soundness/methodology	4

Scale: 1:Lowest Value;6:Highest Value

#### Observations for Author

The authors introduce and detail a tool to support reasoning about failure in the objective to reduce the need infrastructure redundancy. They present data generated by the tool to illustrate its use. They finish the presentation with some extensions and applications for this tool. The paper is quite well presented. It gives first a state of the art with related works. Then it presents a D-PLG. It finishes by giving experimental results. The reviewer is not a specialist of the domain but the paper can be followed and understood quite easily. From my point of

view it seems it can be considered as a M&S approach but it is not completely clear for me what can be the specification (Model) of the on adaptive failure prediction. I would encourage extending the section 3 dedicated to situate the interest of such approach regarding the field of modeling and simulation. For instance this section (3) would gain if having more illustrations and schema (or models).

#### Review #: 5

Abstract and Introduction are adequate?	No
Conclusions/Future Work are convincing?	No
Figures are adequate ? in number and quality	No
Improve critical discussion ? validation	Yes
Improve English?	Yes
Needs comparative evaluation?	Yes
Needs more experimental results?	No
Originality Newness of the ideas expressed	3
Overall Rating Weighted value of above items	3
Paper formatting needs adjustment?	No
Presentation Structure/Length/English	3
References are up-to-date and appropriate?	Yes
Relevance Paper fits one or more of the topic areas?	3
Significance Is the problem worth the given attention?	2
Technical Quality Theoretical soundness/methodology	3
Scale: 1:Lowest Value;6:Highest Value	

#### Observations for Author

1) When reading the abstract one gets an impression that the paper will be about failure prediction. Unfortunately, this is not the case -- the paper only presents a load generation tool "D-PLG". Also, part is Related work is devoted to failure prediction which is not really relevant given the results presented later in the paper which have nothing to do with this. 2) The authors claim that existing approaches to load generation do not emulate the load in a realistic way in a Microsoft active directory environment which is the target for the D-PLG tool. This claim is substantiated but only partially, as far as emulation of the entire software stack / protocol is concerned. However, a more important concern is how to model / simulate the workload in a realistic way using the D-PLG. This is only mentioned in future work which is a little disappointing. 3) Ultimately, the paper presents a technical improvement rather than a scientific contribution. I would expect that a scientific paper would be about one of the problems described in the Future work section, using the tool, rather than about the tool itself which is a means to an end. 4) The experimental results do not seem to prove much other than the tool is efficient and effective. Again, this is not a scientifically interesting result. It would be one if the authors could demonstrate that the proposed tool produces to a more accurate load simulation than existing approaches. D-PLG seems to have potential for this but it needs to be proven. 5) The paper could definitely use more figures illustrating some of the description (for example in section 4.1).