Quantitative Aggregation – An approach for Volume Testing

Author: Ashok Pantham
Deloitte Consulting India Private Limited
Block A, Deloitte Drive, Road No. 2, Hi-tech City Layout,
Hyderabad - 81, India

Author: Jitin Prakash Kulshrestha
Deloitte Consulting India Private Limited
Block A, Deloitte Drive, Road No. 2, Hi-tech City Layout,
Hyderabad - 81, India
Abstract - Time and again we see the importance of Performance testing for any web application. Performance testing is an umbrella which covers many types of testing under it. One important type, which we would like to address in this paper, is Volume Testing. Volume Testing becomes very important to do accurate capacity planning of an application. Any application under test should be able to perform under defined SLAs when tested on a considerable volume. Now question is what is the right Volume of data on which we should test our applications?

In this Paper, we would like to present an approach to weigh all the factors and calculate the right volume for the volume testing of a web application. There are various factors which drive us to calculate the volume for our volume tests like, current volume of Production database, no of users in production for a certain period, usage forecast for near future etc. But do we have a specific method or formula to weigh these factors and calculate right volume of data?

We would like to introduce a new approach, which we call “Quantitative Aggregation”, which guides us in right direction while weighing all these factors. This approach will specify all the important factors required for calculation of Volume for our Volume Testing. It also specify the formula using which each factor can be assigned a weightage and a quantitative aggregation can be performed to come up with the right volume of data.

Why Volume Testing?

“Volume testing” is a term given and described in Glenford Myers’ The Art of Software Testing, 1979. Here’s his definition: “Subjecting the program to heavy volumes of data. The purpose of volume testing is to show that the program cannot handle the volume of data specified in its objectives” – p. 113.

From the definition it is clear that during volume test the application under test is subjected to huge volume of data. This data might be part of database or the file system. Some application might use flat files on file system for operations.
The objective also covers all types of performance testing on top of the volume. This gives a hint of how the application will behave after two years after same amount of transactional data would be generated.

**How do you calculate the required volume?**

Till now, there is no specific formula or method to calculate the right volume on which you have to test your application. Usually People identify the size of this volume on a guess. Some people consider few aspects like future forecast, current volume of production etc. but there is no guideline on how to use them.

Let us first discuss about the aspects that need to be considered before calculating the right volume.

**Current Volume of Production**

If the application is already live in production, the current volume of production plays an important role in base lining the volume at the beginning of Volume calculation. It can be considered as baseline and should be analyzed with historical data.
Historical Data

Historical data should be considered while calculating Volume. It helps us to understand peak periods during a year and also gives an idea of the trend of data accumulation in database of application.

Forecast for near future

Forecast for near future must also be considered for calculation of Volume for Volume testing. Organization generally forecast their growth based on their business plans. It plays an important part in the capacity planning of the application.
**Capacity Planning**

Capacity planning is an important factor while concluding the right data for volume testing. Capacity planning is a broad spectrum which helps in other types of performance testing as well. Capacity planning considers the factors like scaling of hardware and network. Capacity planning to certain extent has to be visualized to decide the volume.

**Other Factors**

There may be other factors that may influence the volume. These factors depend on individual applications as per their functional and business requirements. For applications that have complex functional requirements may have multiple such factors.

**Multiplication factor**

All the above factors that we have seen should be used to drive a multiplication factor. That multiplication factor is used for calculating the volume. It plays an important role in this approach. It is the core of this method using which we take care of all the aspects.
**Multiplication unit**

Multiplication units are the entities that should be multiplied with multiplication factor. All these units once multiplied with multiplication factor are then aggregated to derive the right volume for volume testing. Multiplication units can be the entities which are directly related to transactional data in database of application. Some common Multiplication units can be as follows.

1) Users  
2) Address  
3) Phone no  
4) Email ids  
5) Age  
6) City

\[
\sum \text{Unit 1} \times MF, \text{Unit 2} \times MF, \ldots \ldots \ldots \ldots \text{,Unit Z} \times MF
\]

**Benefits of this concept**

- This approach will require Performance Testers to collect all the right information for specified factors. Based on the information, an exercise to assign weightage to each factor will be done and data volume will be calculated.
- This approach will make a specified framework which will be used by performance testers.
- By using this approach, they will be aware of all the aspects required for Volume testing.
- They will be aware of all the questions that they need to ask business teams about data forecast.
- They will be clear with their objective and can use this approach consistently across different projects.
Conclusion

This technique may take some time initially to do all the brainstorming but it surely reduces and a lot of risks by finding right volume for volume testing. More and more performance issues can be found if we use this method effectively.
People may have their own opinions regarding this concept but they all will agree that this approach gives a methodical shape to the calculation of volume. Let us all help each other in embrace this revolutionary change.

References

1) Performance testing guidance for web applications by J.D. Meier, Carlos Farre, Prashant Bansode, Scott Barber, Dennis Rea
2) www.perftestplus.com
3) The art of software testing by Glenford Myers, 1979