

How To Improve Performance of High Traffic Web Applications

Muraleekrishna Radhakrishnan

MS Computer Science, Texas A & M University, Texas,
USA (2007 pass out)

Abstract- Information Technology uses numerous types of tools for information sharing and gathering of new information. For e.g. Emails, VOIP messages, Websites, Web Applications and so on. Among these tools, web applications are very powerful because web applications can include other tools also. Web applications can include email function. They also can include VOIP interface. Whenever we need processed data from the web, we are going to use web applications. Suppose we want to check the current temperature in your area using web. We just need to type the appropriate web address that points to the web application that retrieves the temperature and the information will be on the screen. But there are many situations where we want the information as quickly as possible. Thus, the performance of web applications is an important factor. I am explaining different ways to improve performance of web applications.

Keywords- Web application, design pattern, layer, database, load balancer, tier

I. INTRODUCTION

Performance of web applications is an important factor when the web applications are used in the enterprise level to share and gather critical information. Enterprise level web applications are usually business applications that include processed data shared among users in a secured way. Usually enterprise level web applications are used by a large number of people from different regions around the globe. The network traffic is a major criterion when this type of application is built.

There are many techniques available to improve the performance of web applications. The techniques differ based on what technology is used to build the applications.

II. TYPES OF WEB APPLICATIONS

While differentiating web applications, there are many ways to do that. Based on the architecture used to build the applications, there are following types of web applications.

1. Standalone applications
2. Distributed applications

III. STANDALONE APPLICATIONS

As the name indicates, standalone applications are applications that usually reside in just one machine. They do not have any dependency to other machines or systems. For e.g. An asp file and a database server reside in the

same machine. But there are CPU usage and memory limitations for standalone applications.

IV. PERFORMANCE OF STANDALONE WEB APPLICATIONS

While looking for ways to improve the performance of standalone web applications, we need to know about the network traffic. If there are many users, then it would be wise to upgrade the client machine in such a way that it can accommodate all user requests. In addition to that the following techniques can be used.

- Use page caching to retrieve the same page when the user requests for it multiple times
- Avoid unwanted database round trips. Sometimes, the data user requested is readily available on the page. Or if the data is too small, we can use hidden variable to store the data. But do not store huge amount of data
- Avoid unwanted loops in the code. Avoid unwanted joins in the SQL stored procedures
- Avoid web server round trip as much as possible by using front end scripts for small operations.

V. DISTRIBUTED APPLICATIONS

Distributed applications are designed in such a way that the applications are built into different layers. The different layers are,

1. Presentation layer
2. Application layer
3. Data layer

Presentation layer comprises of html files, asp files, views etc. and the presentation layer logic to connect to the application layer. The presentation layer logic does the necessary operations to communicate with application layer and the user interface. It accepts user inputs and displays information. Application layer consists of module that facilitates communication between presentation layer and data layer. Usually web services do this task. The data layer has the business logic to do the database operations.

Thus, distributed applications are multi-layered. The purpose of having multiple layers is that it gives the opportunity to distribute the whole solution among different servers. Distributed applications utilize less memory when a single machine is considered. Even though it utilizes more resources, there are many advantages. Multi-layered applications are divided into other categories based on number of layers they have. They are,

1. Two layered applications
2. Three layered applications

Two layered applications have only two layers once the application is built. They have only presentation layer and data layer. The presentation layer directly communicates with the data layer. Three layered applications have application layer which is actually a web service that is a channel to facilitate communication between presentation layer and data layer.

VI. PERFORMANCE OF DISTRIBUTED APPLICATIONS

Distributed applications that have three or more layers have an opportunity to utilize several resources to complete the task. Since they have multiple layers that communicate using web service, the layers can reside in remote servers. And these layers can call web service to perform operations. In this type of architecture, the presentation layer does not communicate directly with the data layer. And also, all the operations defined in the data layer are not available to the presentation layer. Only those operations that are available through web service as web methods will be available for the presentation layer. This gives the opportunity to reuse the same data layer for different presentation layers of other applications. Presentation layer also can call different web services to do the database operations. This gives a lot of flexibility for distributed applications.

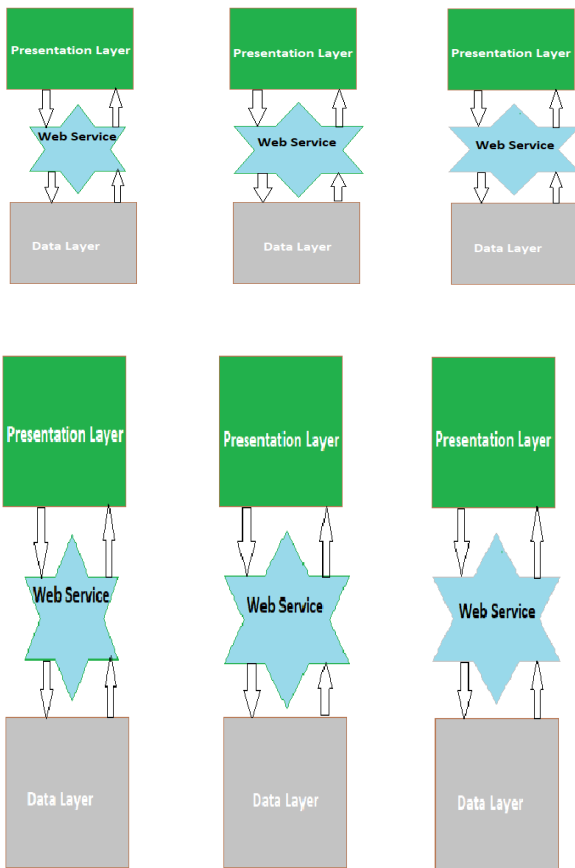


Fig.1: Distributed Applications

As shown above, the presentation layer calls web services whenever there needed to be database operation to be carried out. There are many web servers that have identical copies of presentation layer which communicates with a web service. There are many data layers that reside in multiple application servers which communicates with its own web service. The performance of distributed applications is improved by the following things.

- Compared to standalone applications, distributed applications need only less CPU usage and memory.
- All techniques used in standalone applications to improve performance can be used in distributed applications as well. Distributed applications can make use of page caching. They can also avoid unwanted database and web server round trips.
- Since there are multiple servers to handle user requests, the servers can be redirected by the network based on the server utilization and availability.
- It is easy to decouple data layer and presentation layer by changing the web service.

VII. OTHER TECHNIQUES USED IN HANDLING NETWORK TRAFFIC

Due to high number of user requests, the servers that host the web applications can run out of resources to handle the requests. In this scenario, there are techniques available to redirect the user requests based on server availability. Load balancer is an example of a technique used to distribute user requests among different servers. It is a network device which detects the load on the servers and selects a server that can handle the user request.

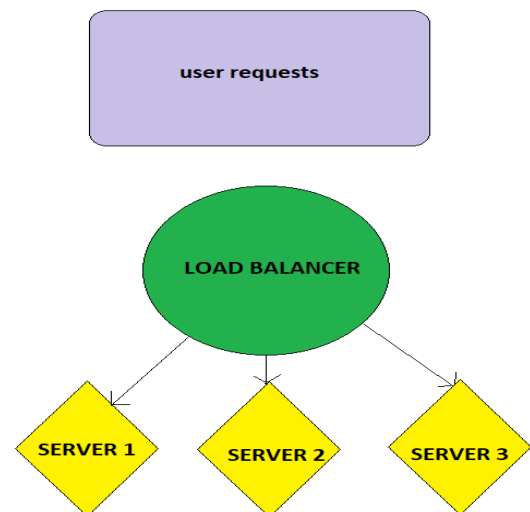


Fig.2: User requests are redirected to different servers by load balancer

VIII. CONCLUSION

As time passes, new technologies are coming into picture and new techniques are available to enhance the performance of applications.

Improving performance of the enterprise level web applications actually increase the productivity of the web users. The main challenge in the Information Technology sector is how to improve performance of applications regardless of number of users. The development work must be carried out in the appropriate way so that the applications will be reliable and efficient.

IX. FUTURE SCOPE OF RESEARCH

The next level of the research is to study the performance comparison of different design patterns in .net. Design pattern is important while an application is built because it determines the way the data flows through the application. Thus, performance varies with the design pattern that was chosen to build the application.

REFERENCES

- [1] Peter Smith, "Professional Website Performance-Optimizing the front and the back end"
- [2] Tony Bourke, "Server Load Balancing"