Introduction to Information Security

Prof. Dilip. Ayyar
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture - 58
Web Application Security

Welcome to module six; here we will deal with web application security. Now in this session we have highlighted some crucial concept or areas from vast top ten and how it matches with sans top twenty five, but web application assessment or audits or security is much more in that. There are several manual methods that you are have to you apart from this ten criteria; lot of tools, lot of processes you will have to call for example, when you are doing a web app assessment, a particular tool - automated tool may not tell you the existence of a captcha or validation of certain fields, so you will have to do that manually. And also there are specific vulnerabilities that will be related to the architecture itself, web server architecture so there is lot of manual process in the assessment of web application.

This will be a starting point was top ten will be a starting point on how you can detect and what are the most common vulnerabilities that will affects the web application. This is not the end of what a web application assessment will be, like I said a lot of manual methods, a lot effort, a lot of tools be it open source, freeware, commercial tools will have to be used for a proper assessment.

You should also know that lot of things comes by experience; it is very easy to run a tool on a website and say that ok these are the lists of vulnerabilities that we have got, but then the real security comes when it is the combination of a web server, your web application, how you server is installed, what kind of security that you put on perimeters side, whether you have web application firewall, whether you have a network firewall or whether you have both a WAF and a network firewall any IDS IPS in places, loss management in place. So many things are process based and manual method based rather than using the tools, so you will have to understand keep that in mind and then go through these slides, and other important factor is no two web application are alike, there the deployment method, the functionality built in to the application, the security features built into the application are may be different so every audit you will have to treat it like a new audit and do it.

Now, this is the first level course of the entire series of five or six percent that we are going to conduct. There will be a separate course on web application security where we will take you step by step on the usage of tools, the usage of manual networks, and how you do a assessment on the web application. This will give you only a over view or rather understanding of the theoretical aspects of different components that constitutes web application security. With this knowledge, I am sure you can attempt web application audits or you can try out different kind of vulnerabilities, there are lots of free sites available. So if you Google that you want to test a vulnerable applications there are many which are on line there are many where you can download to VM in installing the system for learning purposes, so you can try it out with that, but please do not try to exploit or hack into a commercial site or any government site, it is illegal even for the purpose of learning. So it is better that you download the Vms, try it on your own machines see how each vulnerabilities is detected, you will get because in web application the more you work on vulnerable apps the better you will be in terms of detecting it and securing it.

So you need to have a very balanced approach when you are doing this, even for applications which are posted which are commercially, but as we go through the slides I will also relate some of the experiences or some of the issues that I encountered during such audits. What you have to understand this do it ethically that the very thin line or the very thin demarcation between use and abuse. So you try to use a knowledge for a put it good use and last abuse systems of others without their permission. With that we will start module six.

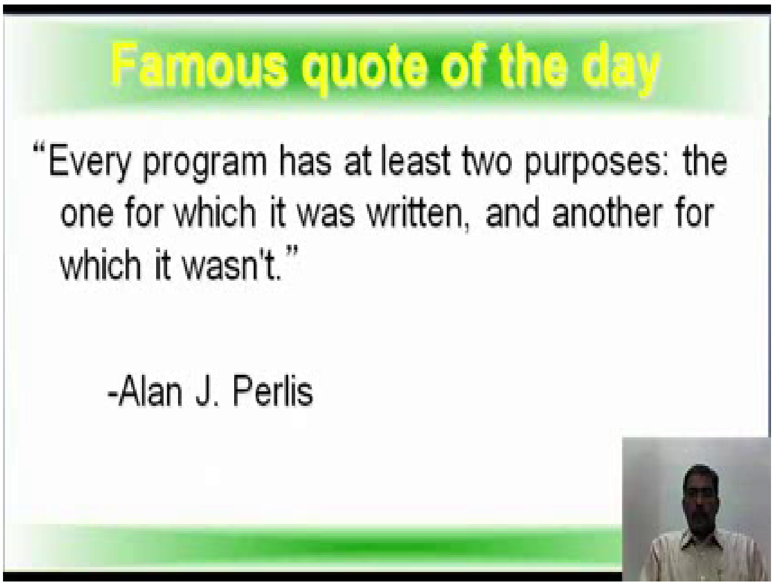
(Refer Slide Time: 05:48)

Goals of this session

- Build security awareness for web applications
- Get to know attack methods of hackers
- Learn ways to discover security vulnerabilities
- Learn the basics of secure web development

What are the goals of this session; one is to build security awareness for web applications. To get to know the attack methods of hackers; learn ways to discover security vulnerabilities; and to learn the basics of secure web development.

(Refer Slide Time: 06:10)



Famous quote of the day

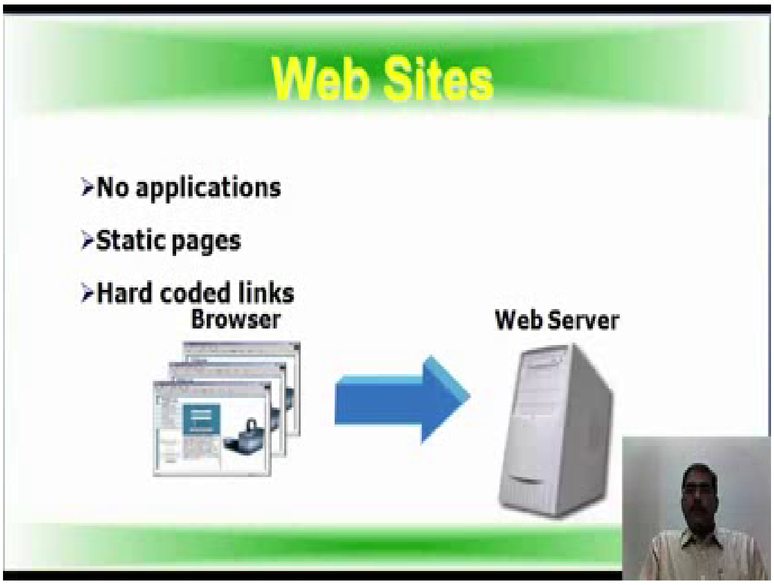
"Every program has at least two purposes: the one for which it was written, and another for which it wasn't."

-Alan J. Perlis

It the very famous quote by Alan J Perlis, every program has at least two purposes: the one for which it was written, and another for which it wasn't. Now big part of web applications security involves attempts to forces an application to function in a way, it was not intended to function. So Alan J Perlis was a computer scientist known for his pioneering work in programming languages; and the first recipient of the Turing award, which is equivalent to the noble prize for computing. To emphasize this further let us take an example of banking applications. There are several fields in the banking applications.

Let us take the specific example of a customer data, where the name, the date of birth, your PAN number, the address all the other details are entered into, your phone number. Now what kind of tests will be conducted on this kind of fields, it is the simple text fields right, but even then it was written for accepting these fields, but what if somebody injects a script one of those field. What is somebody enters numerical values in a text field, how the system will behave, when it is not supposed to. How your system will to behave with somebody injects the cross side of introduces of cross side scripting vulnerability. Now all of this is something with the called as negative testing or a testing for which or purpose for which it was not written for. Now all this happens because of faulty coding, lack of knowledge of security from the development team or lack of coordination between the information security team and the development team. It can also be because of faulty configuration in the server, in the web server, so it could be any of those.

(Refer Slide Time: 08:23)

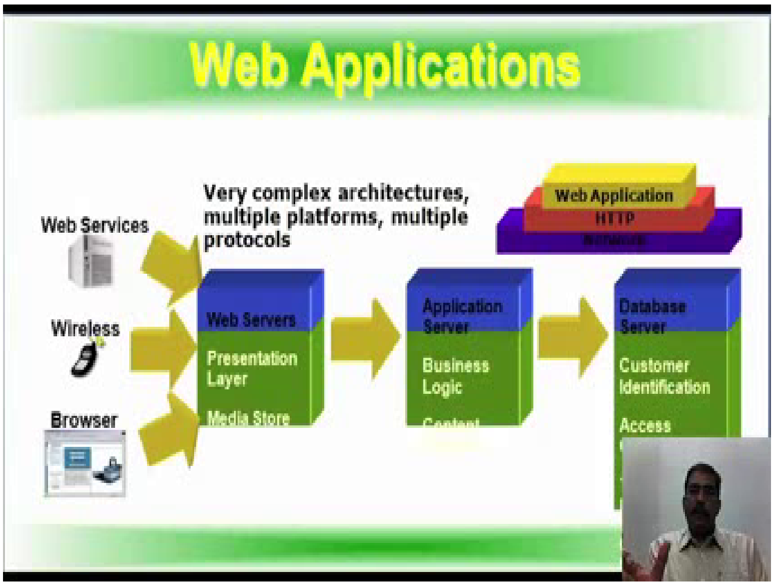


Web Sites

- No applications
- Static pages
- Hard coded links

So we will be talking about web application, websites. So what is a websites, it is not an application, they are static pages may be if you take the websites of a company, it may have a home, about us page, contact us pages, products and services pages, clients page, so through are static pages information does not change frequently. There are hard links or hard coded links to the browsers, and the web application or website sits on a web server. What are the kind of web server IAAS for the Microsoft is a web server; Apache is a web server, Comcat so they are so many engine x so there are so many web servers also available. So in a website, the website is created and it is loaded on a web server.

(Refer Slide Time: 09:17)

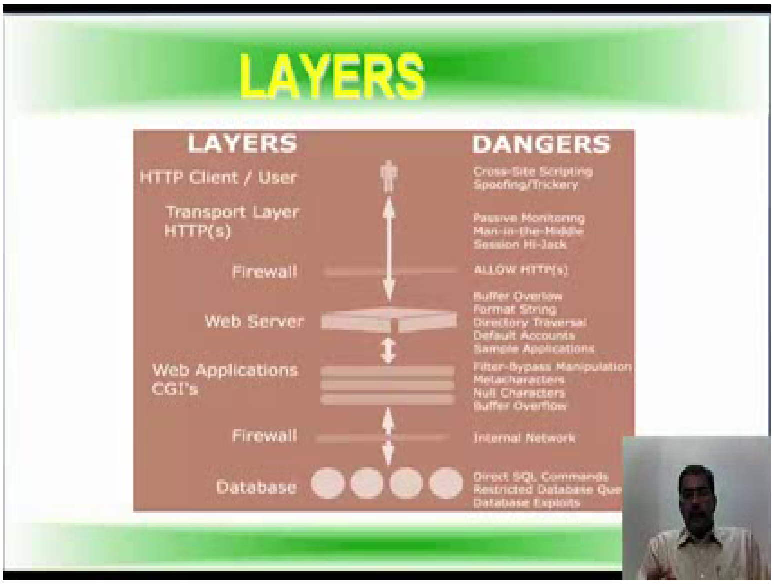


Web Applications

Very complex architectures, multiple platforms, multiple protocols

Now, we know that website is static, static content is there. Now when you talk about web applications, it got very complex architectures, multiple platforms with multiple protocols enabled. So if you see from the left side on wards there are web services, which get connected to the web servers wireless that is your mobile apps server and then from your desktops your browser is there. There are three layers that one is the web server layer, the presentation layer and the media store. From there it goes to the application server, where the business logic is written and then the contents are puts. From there, it is passes onto the database server where the customer identification, access control, the transaction information, the core business data all are loaded into that. So what we have to understand that is a web application which works on a http protocol and then there is a network from where different kinds of requests will come in; whether it is web services, whether it is a wireless whether it is through a desktop, whether it is through ipad, so whatever it is so different layers are there in the web applications.

(Refer Slide Time: 10:37)

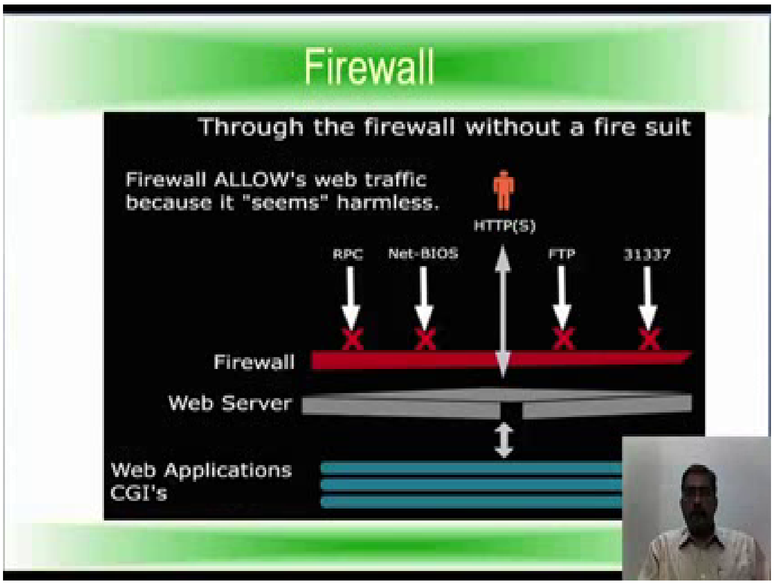


LAYERS

LAYERS	DANGERS
HTTP Client / User	Cross-Site Scripting Spoofing/Trickery
Transport Layer HTTP(s)	Passive Monitoring Man-in-the-Middle Session Hi-Jack
Firewall	ALLOW HTTP(s)
Web Server	Buffer Overlow Format String Directory Traversal Default Accounts Sample Applications
Web Applications CGI's	Filter-Bypass Manipulation Metacharacters Null Characters Buffer Overflow
Firewall	Internal Network
Database	Direct SQL Commands Restricted Database Que Database Exploits

We were talking about different layers. Now if you look at this you watch the center one that is a client http client or a user. The information goes through the transport layer, it can be https traffic; it actually passes through a fire wall where the web server is installed. After that the information passes through the web server; from there, the information again requests pass through the web application CGI then to another firewall and then finally to the data base. Now if you look at the right hand side, the dangers, we will see in the client place, you will find excesses, spoofing, trickery happening at that layer then passive monitoring or man in the middle attack or the session hijacking will happen in the transport layer. Your firewall allows only http or https then the web server, you have buffer over flows; format string attacks, directory traversal, defaults accounts sample applications now all those dangers will be there in the web server in web applications CGI, we have filter by first matriculation you have meta characters filtering, null characters buffer over flows all those will come under web application CGIs. Then there is another firewall which segregates the external network from the internal network or the web application from the data base. Then in database itself you have SQL injections database exploits, restricted database queries; so all those will affect that particular layer.

(Refer Slide Time: 12:31)



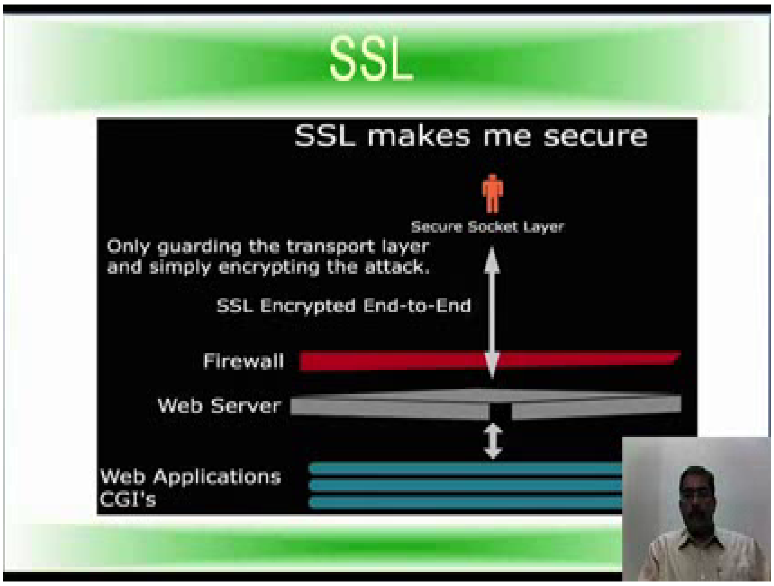
Firewall

Through the firewall without a fire suit

Firewall ALLOW's web traffic because it "seems" harmless.

Even though, we have seen the contexts of the firewall in domain port. This is specific to the web applications. Now through the firewall without the fire suit meaning you just walk through the firewall without the suit, we protecting suit. Firewall allows web traffic because it seems harmless so whether it is your http or https traffic it allows. The other services, if you can see red bar RPC, your net bios, FTP, port 31337 all are block. So http or https allow to the firewall, so you are able to walk through the firewall. Then if the web server the request and then to the application CGI's and so on it goes down, so the firewall in this contexts is configured in such a way that only http or https traffic is allowed.

(Refer Slide Time: 13:32)



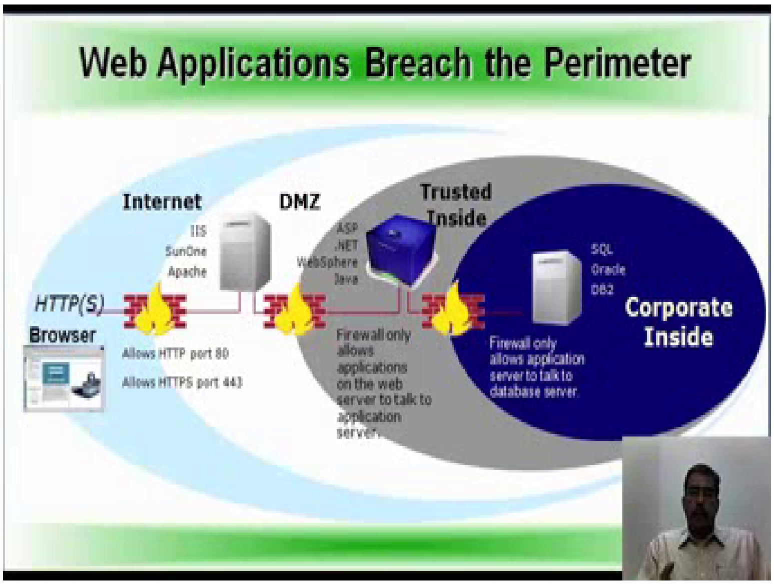
SSL

SSL makes me secure

Only guarding the transport layer and simply encrypting the attack.

Then we have SSL - secure socket layer. So we think SSL makes you secure, so what it does only guarding a transport layer and simply encrypting attack. So if I am formatting a particular request to be sent to the server by using SSL I am encrypting the attack and setting it through, now it will pass through the firewall because port 80 and 443 are allowed, it does not read that contents of that particular request then it goes to the web server and web applications, internal firewall and database.

(Refer Slide Time: 14:10)



Web Applications Breach the Perimeter

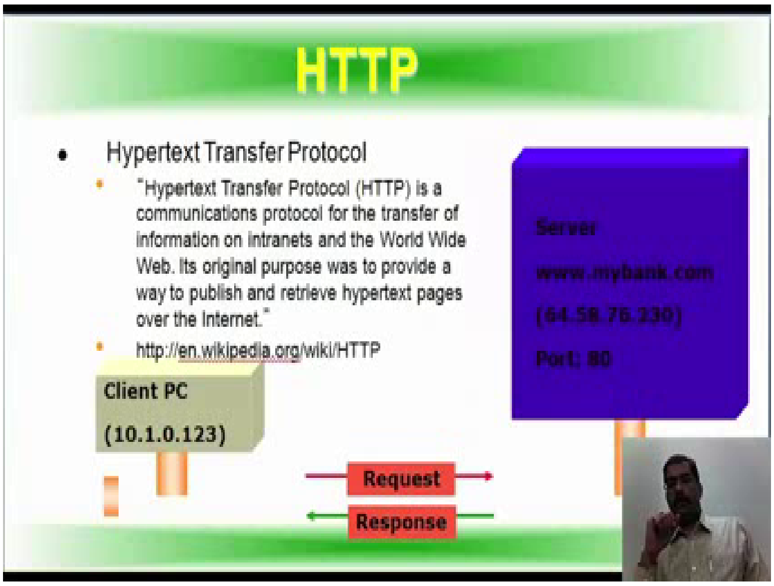
Internet
DMZ
Trusted Inside
Corporate Inside
IIS
SunOne
Apache
ASP
.NET
WebSphere
Java
SQL
Oracle
DB2
HTTP(S)
Browser
Allows HTTP port 80
Allows HTTPS port 443
Firewall only allows applications on the web server to talk to application server.
Firewall only allows application server to talk to database server.

Now, take the example here, web applications breach the perimeters. You have a http a browser from which http or https request is allowed; http is port 80; https port 443 through the firewall it goes to web server, it can be a Apache, IIS or SunOne; from there, it goes to the dematerialized zone where the web server are installed. From there, it goes to another firewall to the trusted inside, so firewalls allow only applications on the web server to the talk to the application server, so the web server and applications communication happens through this firewall and it allow only that traffic. Then the firewall there is another firewall if you see the dark blue ring that firewall only allow applications server to talk to the database server, so the data base server can be SQL, Oracle, DB2 whatever it is, and it is a cooperate environment, if it is installed within the cooperate environment.

Now, the latest method of doing is you have firewall which filters based on these ports. You also have a web application firewall; a web application firewall actually reads contents require that is convenient and it is decides whether to allow or deny that particular request. Now, this firewall is network based that one depicted; the web application firewall, I will bring the mouse over to the place, this generally installed here or here so that here web request are analyzed based on the character based on different parameters of whether it is OAS or PCA D or whatever criteria and set or your own criteria and then it decide whether to allow the traffic further inside or drop the packet right away.

So, web application firewall should actually form a part of this diagram, but since it is only gaining popularity of-late; many organizations do not use web application firewall, but your payment gateway systems, the banks, your insurances companies all of them use web application firewall because every request is analyzed and then the web application firewall decides whether to allow it or deny it why this organization use it because the value of data is very critical; loss of that data will cause havoc, there is a lot of loss of reputation, the customers can file a complaint against the bank, they can file the loss suit against the bank. So a web application firewall forms the crucial part of this environment.

(Refer Slide Time: 17:17)



HTTP

- Hypertext Transfer Protocol
  - "Hypertext Transfer Protocol (HTTP) is a communications protocol for the transfer of information on intranets and the World Wide Web. Its original purpose was to provide a way to publish and retrieve hypertext pages over the Internet."
  - http://en.wikipedia.org/wiki/HTTP

Client PC
(10.1.0.123)

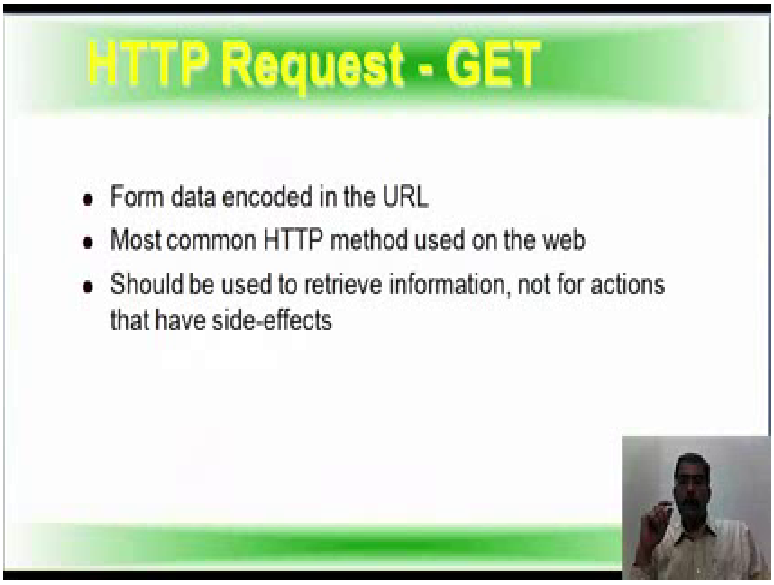
Server
www.mybank.com
(64.58.76.230)
Port: 80

Request

Response

HTTP which is called hypertext transfer protocol; http is a communication protocol for the transfer of information on the intranets and the World Wide Web. Its original purpose was to provide a way to publish and receive hypertext pages over the internet. This definition is as per Wikipedia. Now if you see an example below, there a client PC, there is a request sent to my bank dot com and the response come all through port 80, http is also called a state less protocol, because each request is process independently. Without any knowledge of the request that came before it, so http is also called the state less protocol.

(Refer Slide Time: 18:10)

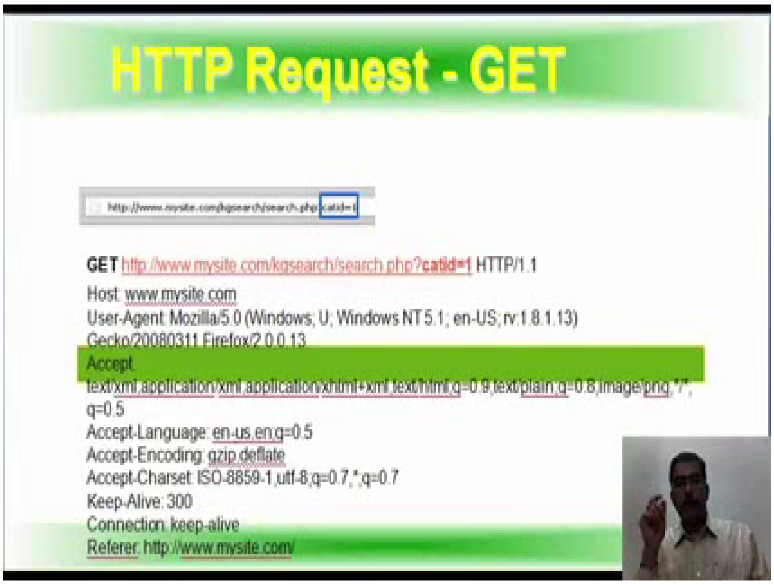


HTTP Request - GET

- Form data encoded in the URL
- Most common HTTP method used on the web
- Should be used to retrieve information, not for actions that have side-effects

Then we come to http request that is a GET request, it is form data encoded in the url. The most common method used on the web should be used to retrieve information not for actions that have side effects. For example, can you send a data with a username and password in a GET request, no, there is something called representational state transfer also, I suggest that you read about it which is rest - r e s t. It is a term coined by Roy fielding in his PhD disseration, to describe the architectural style of the network system. So to understand this better you can actually go and see what is REST- r e s t and learn about it.

(Refer Slide Time: 19:03)



HTTP Request - GET

```
http://www.mysite.com/kgsearch/search.php?catid=1

GET http://www.mysite.com/kgsearch/search.php?catid=1 HTTP/1.1
Host: www.mysite.com
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.8.1.13)
Gecko/20080311 Firefox/2.0.0.13
Accept:
text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8,image/png,*/*;
q=0.5
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
Referer: http://www.mysite.com/
```

This is a simple get request like we see there is some mysite dot com, from search search p h p question mark category i d equal to 1, so there is some form encoded or there is something encoded in the url. Now this is how the responses when you actually to request responses, analysis, there is a host, a user agent, accept is there, http request should be accepted by the server, the language, the keep alive connection from where it is who has refer this particular site so the referrer, we will talk about the referrer and the user agent update later on.