CDN (Content Delivery Network)

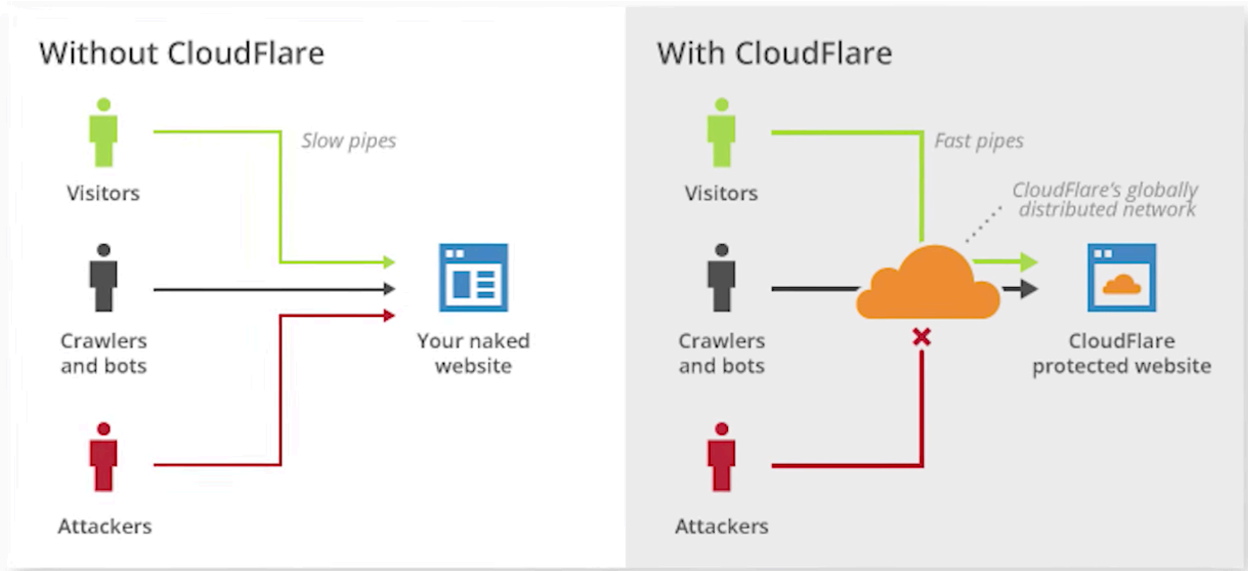
What is CDN:

- A content delivery network (CDN) is a system of distributed servers (network) that deliver pages and other Web content to a user, based on the geographic locations of the user, the origin of the webpage and the content delivery server.

- This service is effective in speeding the delivery of content of websites with high traffic and websites that have global reach. The closer the CDN server is to the user geographically, the faster the content will be delivered to the user. CDNs also provide protection from large surges in traffic.

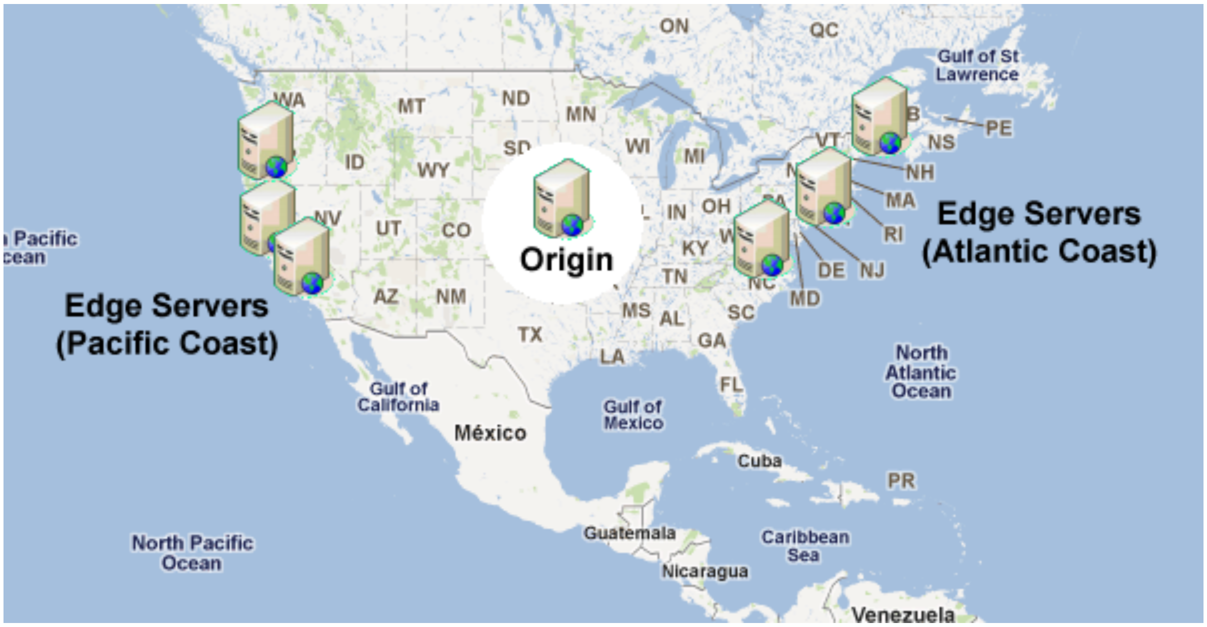
How CDNs work:

- Physics determines how fast one computer can contact another over physical connection, and so attempting to access a server in China from a computer in the United States will take longer than trying to access a U.S. server from within the U.S. To improve user experience and lower transmission costs, large companies set up servers with copies of data in strategic geographic locations around the world. This is called a CDN, and these servers are called edge servers, as they are closest on the company's network to the end-user.



DNS Resolution:

- When the browser makes a DNS request for a domain name that is handled by a CDN, there is a slightly different process than with small, one-IP sites. The server handling DNS requests for the domain name looks at the incoming request to determine the best set of servers to handle it. At its simplest, the DNS server does a geographic lookup based on the DNS resolver's IP address and then returns an IP address for an edge server that is physically closest to that area. So if I'm making a request and the DNS resolver I'm routed to is Virginia, I'll be given an IP address for a server on the East coast; if I make the same request through a DNS resolver in California, I'll be given an IP address for a server on the West coast. You may not end up with a DNS resolver in the same geographic location from where you're making the request.

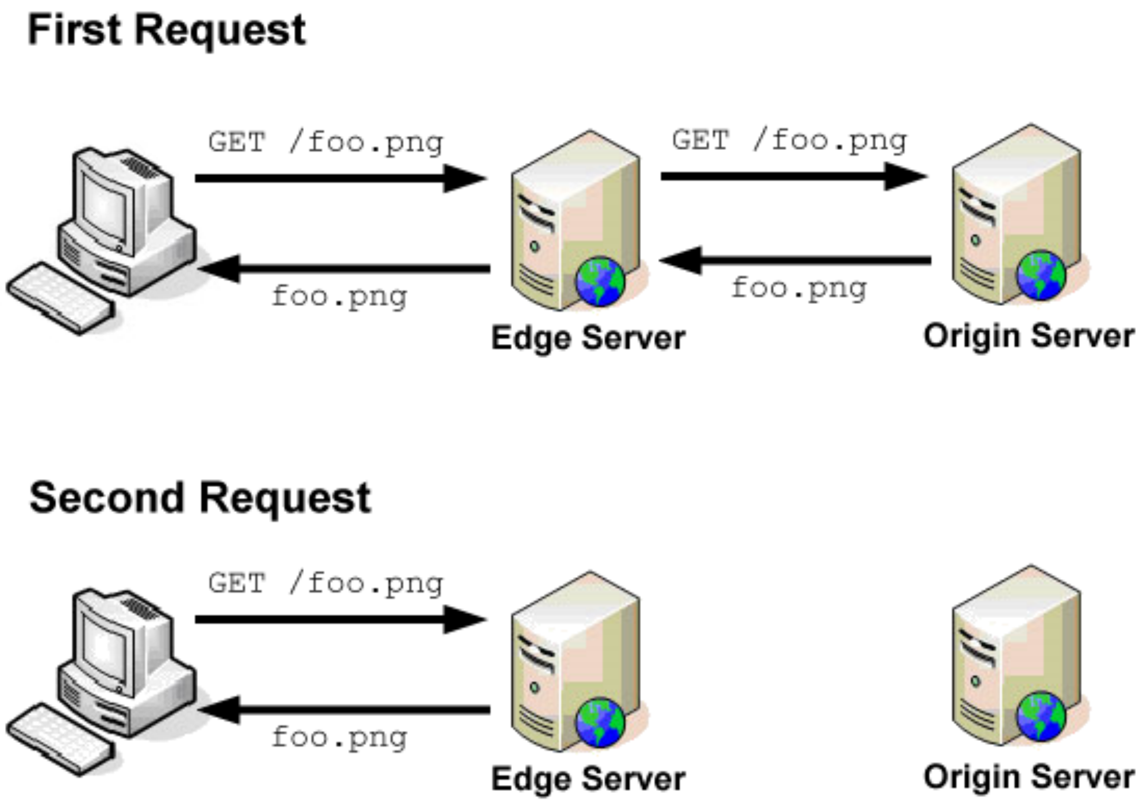


That's the first step of the process: getting the request to the closest server possible. Keep in mind that companies may optimize their CDNs in other ways as well, for instance, redirecting to a server that is cheaper to run or one that is sitting idle while another is almost at capacity. In any case, the CDN smartly returns the best possible IP address to handle the request.

Accessing Content:

- Edge servers are proxy caches that work in a manner similar to the browser caches. When a request comes into an edge server, it first checks the cache to see if the content is present. The cache key is the entire URL including query string (just like in a browser). If the content is in cache and the cache entry hasn't expired, then the content is served directly from the edge server.

- If, on the other hand, the content is not in the cache or the cache entry has expired, then the edge server makes a request to the origin server to retrieve the information. The origin server is the source of truth for content and is capable of serving all of the content that is available on the CDN. When the edge server receives the response from the origin server, it stores the content in cache based on the HTTP headers of the response.



- Yahoo! created and open sourced the Apache Traffic Server, which is what Yahoo! uses in its CDN for managing this traffic. Reading through the Traffic Server documentation is highly recommended if you'd like to learn more about how cache proxies work.

EXAMPLE:

For example, Yahoo! serves the YUI library files off of its CDN using a tool called the combo handler. The combo handler takes a request whose query string contains filenames and concatenates the files into a single response. Here's a sample URL:

http://yui.yahooapis.com/combo?3.4.1/build/yui-base/yui-base-min.js& -
038;3.4.1/build/array-extras/array-extras-min.js

The domain yui.yahooapis.com is part of the Yahoo! CDN and will redirect you to the closest edge server based on your location. This particular request combines two files, yui-base-min.js and array-extras-min.js, into a single response. The logic to perform this concatenation doesn't exist on the edge servers, it only exists on the origin server. So if an edge server receives this request and has no content, a request is made to the origin server to retrieve the content. The origin server is running the proprietary combo handler (specified by /combo? in the URL) and so it combines the files and returns the result to the edge server. The edge server can then serve up the appropriate content.

What does static mean?

- frequently get confused looks when I describe systems similar to the combo handler. There is a misconception that CDNs act like FTP repositories, where you simply upload static files so that others can retrieve them. I hope that it's clear from the previous section that this is not the case. An edge server is a proxy, the origin server is the one that tells the edge server exactly what content should be returned for a particular request. The origin server may be running Java, Ruby, Node.js, or any other type of web server and, therefore, can do anything it wants. The edge server does nothing but make requests and serve content. So, the YUI combo handler exists only on the origin server and not on the edge servers.

- If that's the case, why not serve everything from the CDN? The CDN is a cache, meaning that is has value when it can serve data directly and not need to contact the origin server. If an edge server needs to make a request to the origin server for every request, then it has no value (and in fact, costs more than just making the request to the origin server itself).

- The reason JavaScript, CSS, images, Flash, audio, and video are frequently served from CDNs is precisely because they don't change that frequently. That means not only will the same user receive content from cache, but all users will receive the same data from cache. Once the cache is primed with content, all users benefit. A site's homepage is a poor candidate for edge caching because it's frequently customized per user and needs to be updated several times throughout the day.

Benefits:

❖ Performance

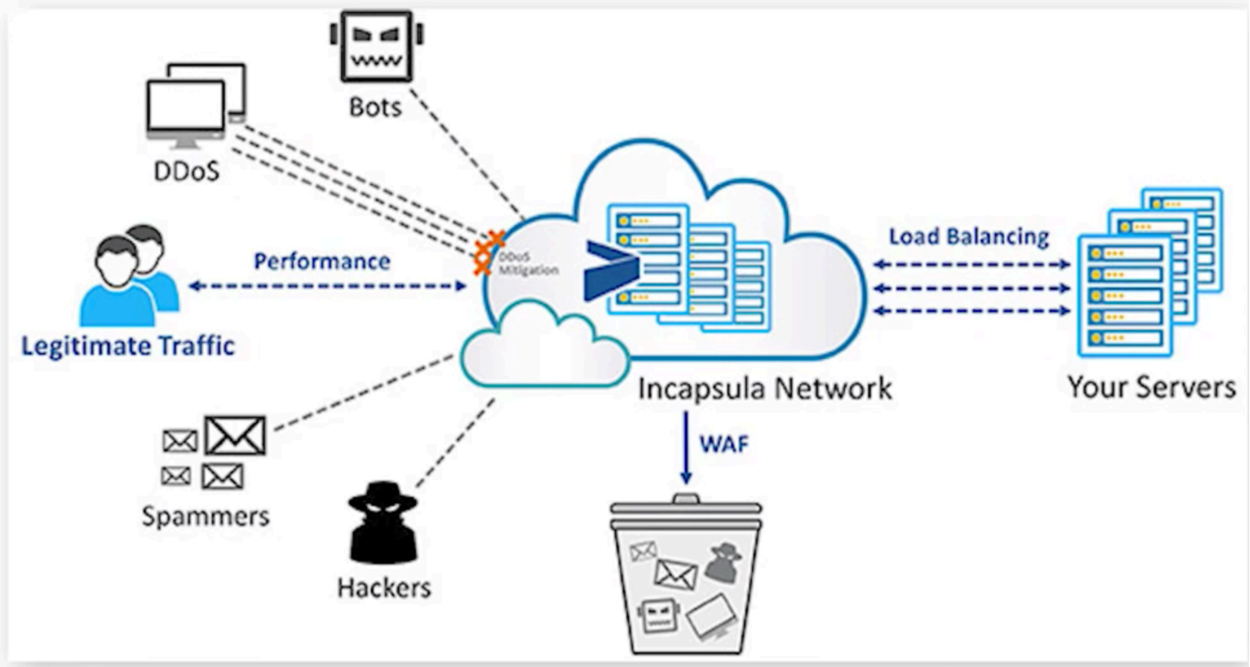
- CDN basically take the snapshot of your website and if someone or a lot of people go to your website, CDN cache the content of the actual website and push it to different servers and tell that person that Hey, there is nothing changed in the website and serve the content to them.

- Let's suppose you have visitors all over the world and your website is hosted in western US. If you are using CDN then the visitor will reach to the nearest CDN and CDN will serve them the cache data. That will also enhance the performance of your servers.

❖ Security

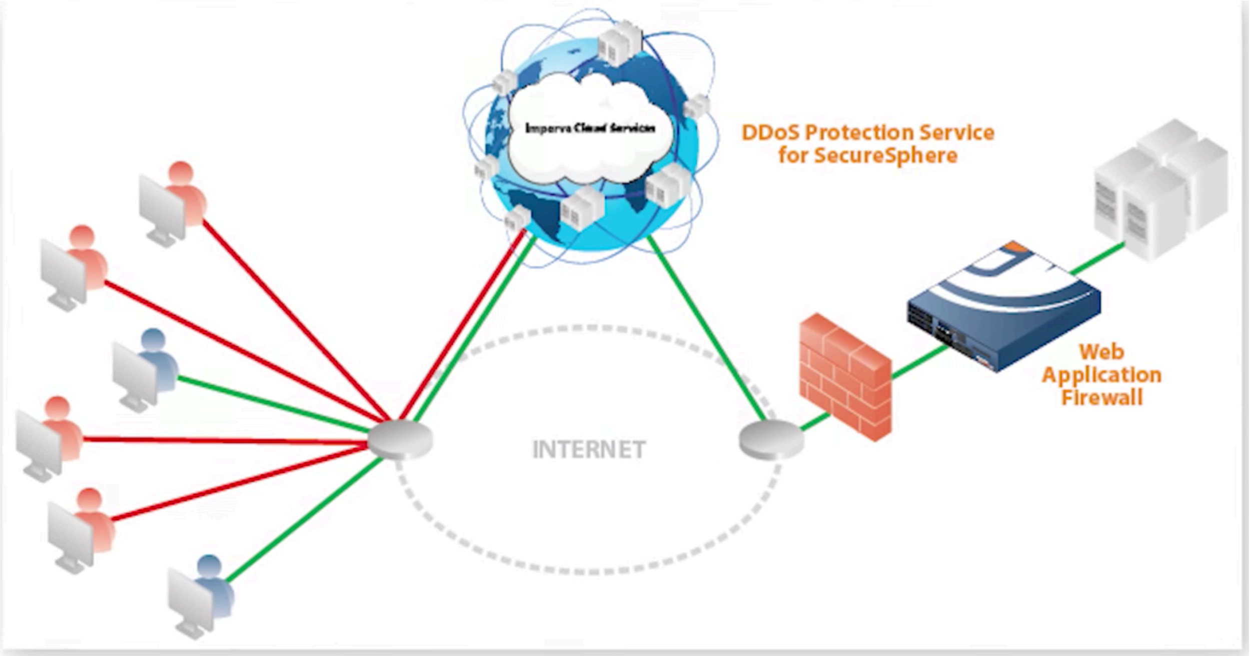
➢ WAF (web application firewall):

With firewall CDN provides a buffer between the person and your website. you can filter out malicious request ever reaching your website in the first place. It detects virus in the website. So, it basically protects incoming and outgoing data.



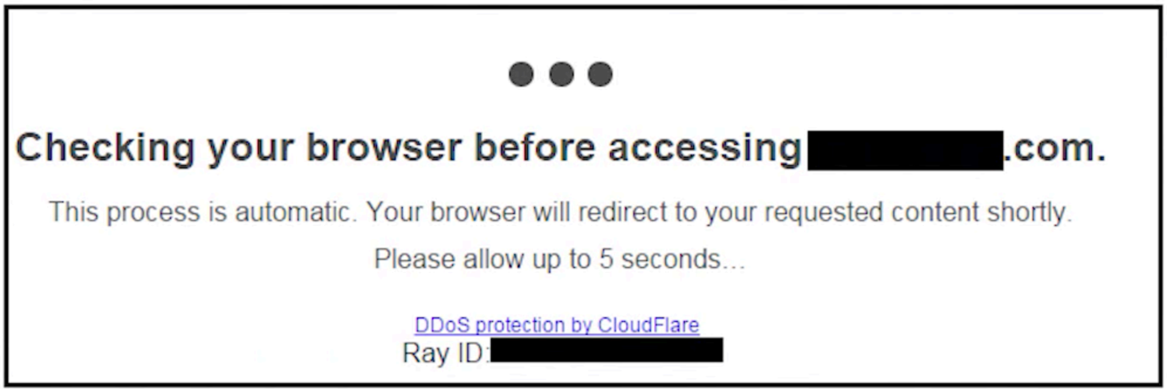
- **D-dos Protection:**

  CDN do offer that protection. what happens is if someone sends a ton of request to the website, so it's sending it to the CDN and not to your actual website so it could be filter out.



Cloudflare

- How Cloud flare shows a message on the browser screen, meanwhile it filters out all the malicious request, bots, viruses.



Checking your browser before accessing [redacted].com.

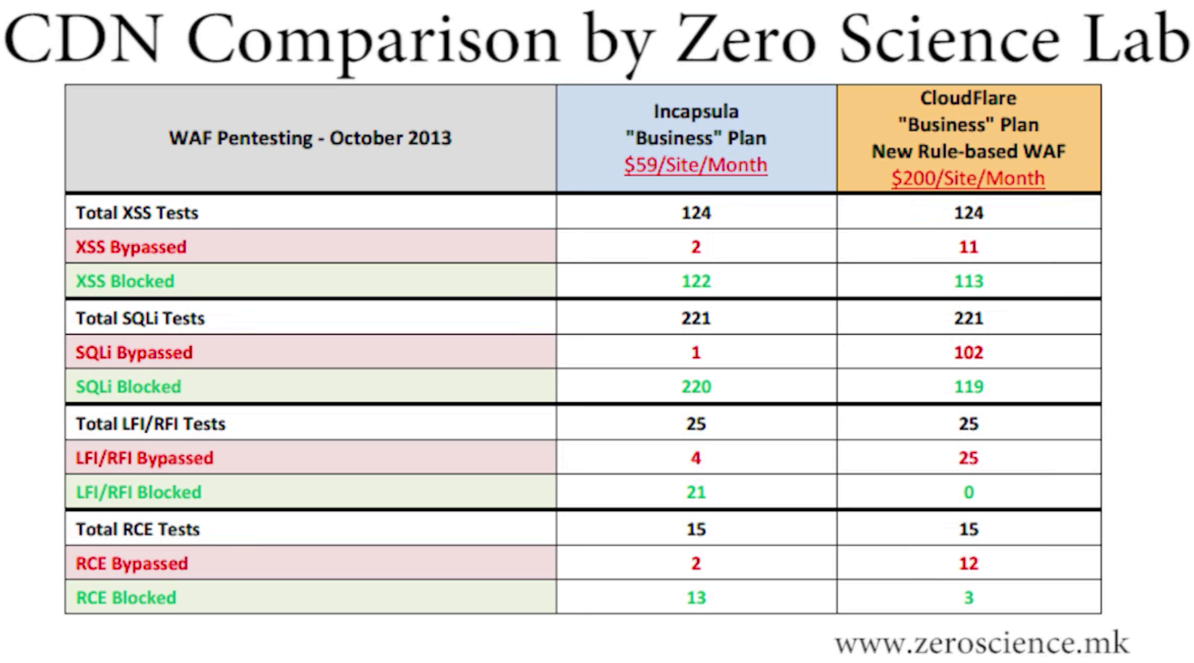
This process is automatic. Your browser will redirect to your requested content shortly.

Please allow up to 5 seconds...

DDoS protection by CloudFlare

Ray ID:[redacted]

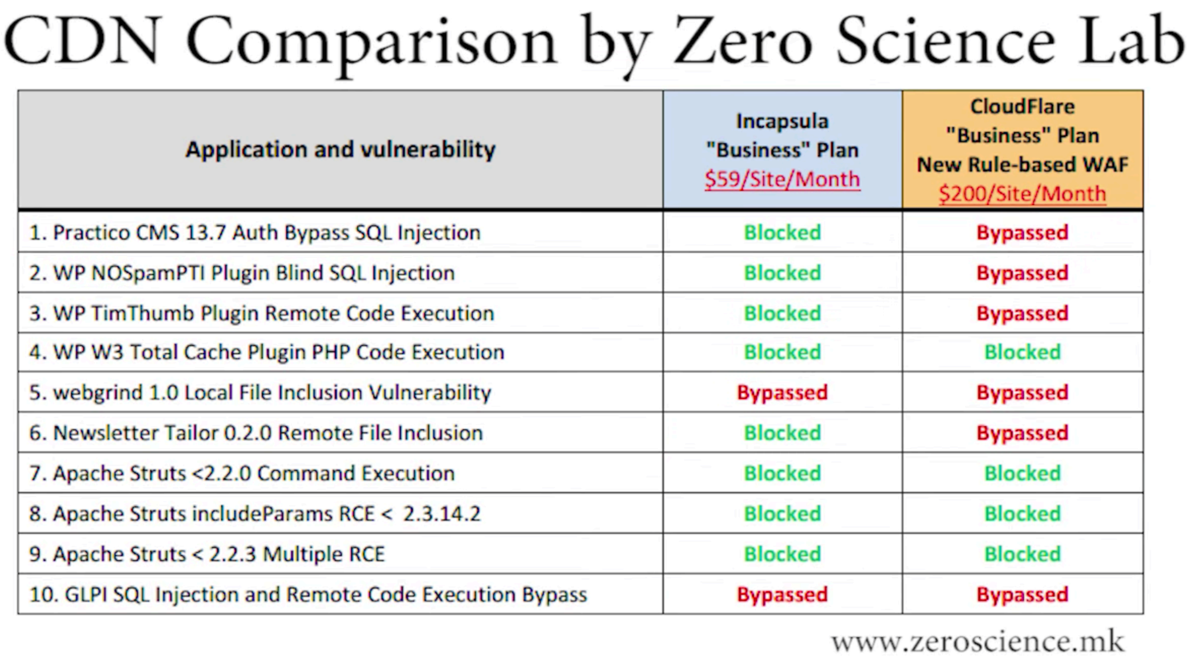
CDN Comparison



CDN Comparison by Zero Science Lab

WAF Pentesting - October 2013	Incapsula "Business" Plan $59/Site/Month	CloudFlare "Business" Plan New Rule-based WAF $200/Site/Month
Total XSS Tests	124	124
XSS Bypassed	2	11
XSS Blocked	122	113
Total SQLi Tests	221	221
SQLi Bypassed	1	102
SQLi Blocked	220	119
Total LFI/RFI Tests	25	25
LFI/RFI Bypassed	4	25
LFI/RFI Blocked	21	0
Total RCE Tests	15	15
RCE Bypassed	2	12
RCE Blocked	13	3

www.zeroscience.mk



CDN Comparison by Zero Science Lab

Application and vulnerability	Incapsula "Business" Plan $59/Site/Month	CloudFlare "Business" Plan New Rule-based WAF $200/Site/Month
1. Practico CMS 13.7 Auth Bypass SQL Injection	Blocked	Bypassed
2. WP NOSpamPTI Plugin Blind SQL Injection	Blocked	Bypassed
3. WP TimThumb Plugin Remote Code Execution	Blocked	Bypassed
4. WP W3 Total Cache Plugin PHP Code Execution	Blocked	Blocked
5. webgrind 1.0 Local File Inclusion Vulnerability	Bypassed	Bypassed
6. Newsletter Tailor 0.2.0 Remote File Inclusion	Blocked	Bypassed
7. Apache Struts <2.2.0 Command Execution	Blocked	Blocked
8. Apache Struts includeParams RCE < 2.3.14.2	Blocked	Blocked
9. Apache Struts < 2.2.3 Multiple RCE	Blocked	Blocked
10. GLPI SQL Injection and Remote Code Execution Bypass	Bypassed	Bypassed

www.zeroscience.mk

Conclusion

- CDNs are an important part of today's Internet, and they're only going to become more important as time goes on. Even now, companies are hard at work trying to figure out ways to move more functionality to edge servers in order to provide users with the fastest possible experience. This includes a technique called Edge Side Includes (ESI) which is designed to serve partial pages from cache. A good understanding of CDNs and how they work is key to unlocking greater performance benefits for users.

References:

https://www.youtube.com/watch?v=kIHHxTWGL8w
https://www.youtube.com/watch?v=P67qQNY5FJI
https://www.nczonline.net/blog/2011/11/29/how-content-delivery-networks-cdns-work/