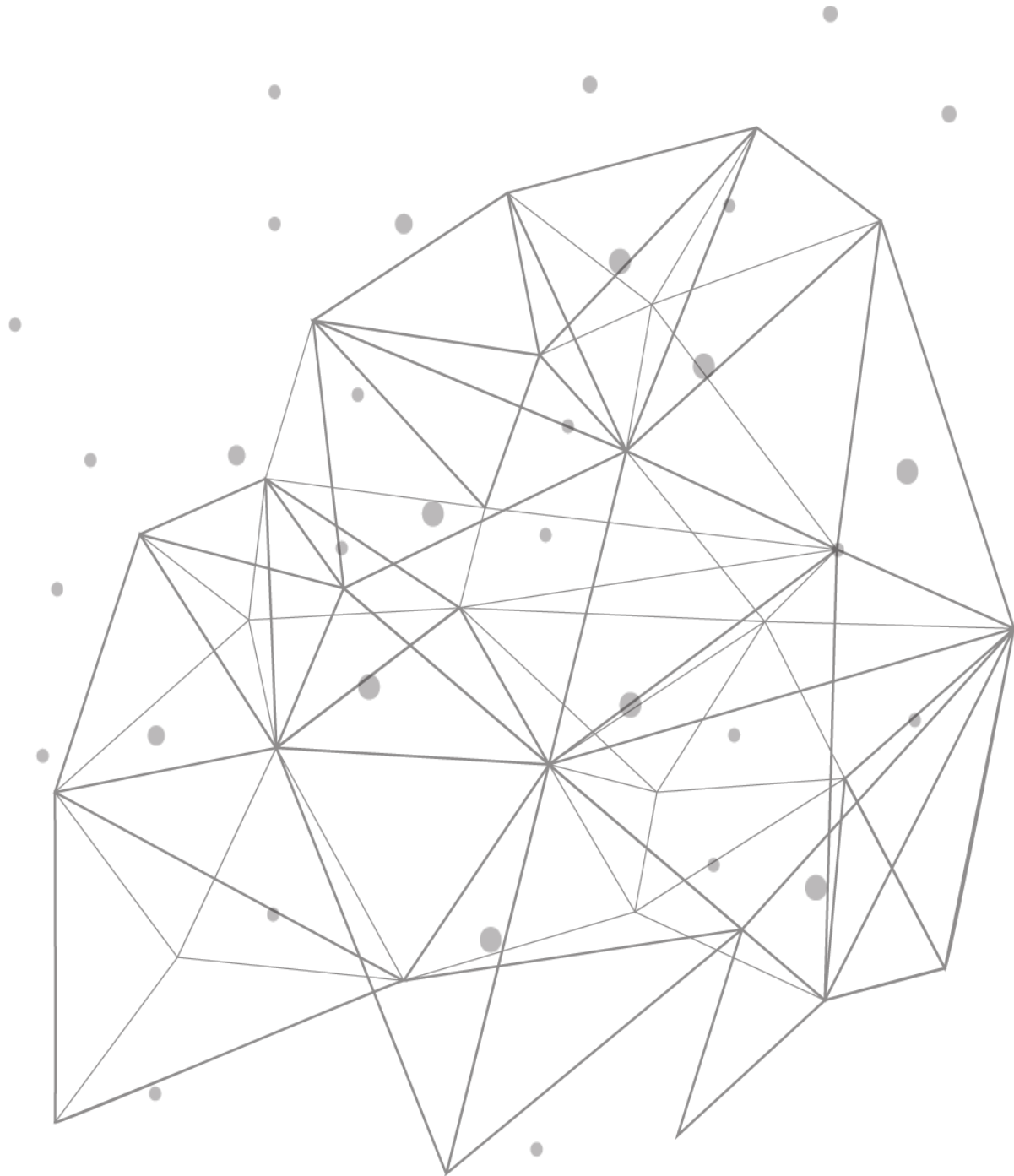


TCPWave IPAM

Cache Hit Ratio



Introduction

Latency is one of the significant challenges for today's network teams, and it dramatically affects clients' experience from an efficiency standpoint and by lowering client satisfaction levels and revenue. One of the core reasons for latency is an inefficient cache policy that minimizes the number of cache hits and maximizes cache misses leading to a lower cache hit ratio.

Cache

A cache is a high-speed memory reserved storage location that temporarily saves data or content from a web page. It helps a web page load faster for a better client experience.

DNS Cache

A DNS cache, aka a DNS resolver cache, is a provisional database supported by the operating system that includes records of all the latest visits to internet domains. In other words, it is a memory of recent DNS lookups that the system rapidly refers to.

Cache Hit Ratio

The effectiveness of a cache is measured through a parameter known as the Cache Hit Ratio, and it is a measurement of how many requests a cache can fill successfully.

Measurement

The following is the formula for a cache hit ratio:

Cache hit ratio = $\left[\frac{\text{cache hits}}{\text{cache hits} + \text{cache misses}} \right] \times 100 \%$

Example:

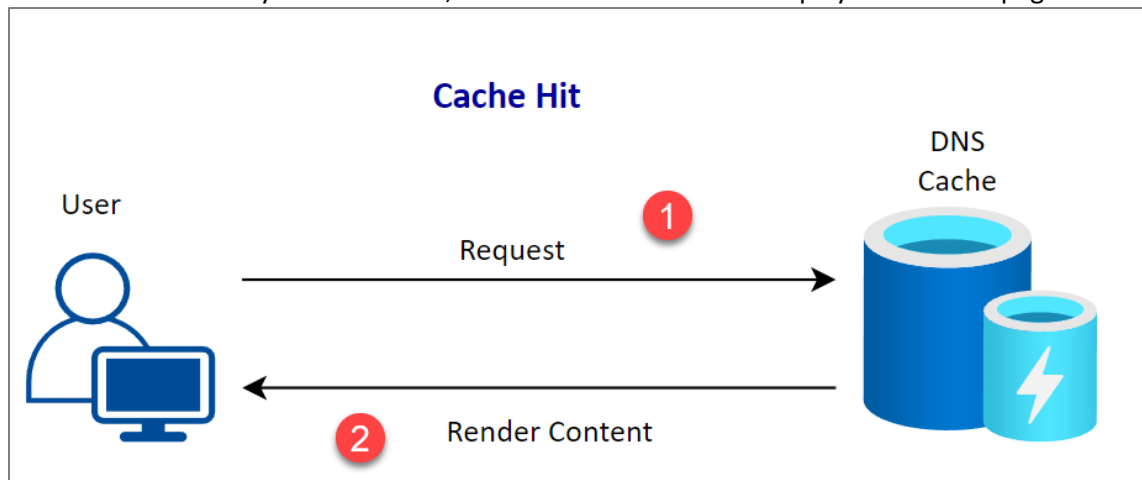
If you have 46 cache hits (requests) and 14 misses, then the cache hit ratio is:

$= \left[\frac{46}{46+14} \right] \times 100 = 76.667\%$

Hence the cache hit ratio is 76.667%. It means that the requested object is served from the cache 76.667% of the time and must be retrieved from the origin server 23.333% of the time.

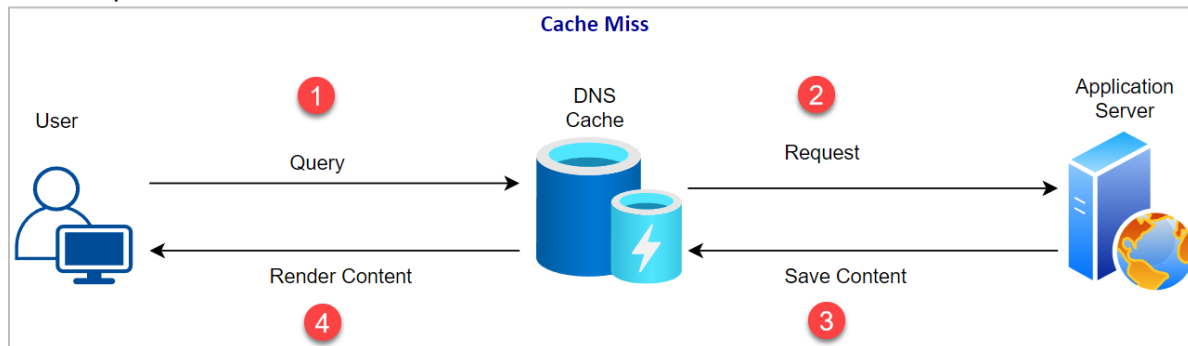
Cache Hit

The cache successfully retrieves data, and content saved to it is displayed on a web page.



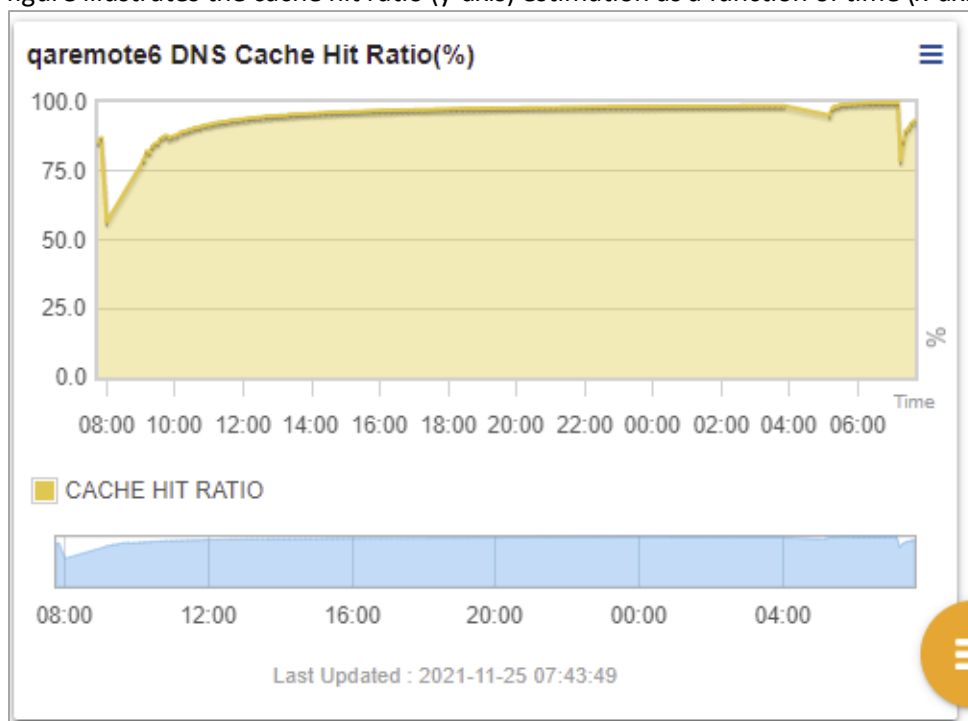
Cache Miss

When the site requests the content from the cache, but after searching for it, if it determines that content was not saved, the cache reserves the content from the application server, so it's accessible when requested the next time.



TCPWave's Cache Hit Ratio Statistics

Monitoring helps the enterprise to gain information about how servers and other devices such as switches, routers, etc., are handling the load generated against the web application. The Domain Name System monitoring is incredibly critical whether the enterprises manage one or many websites, and it helps the enterprises keep the services secure and safe. Additionally, it improves the browsing experience, reducing costs in terms of bandwidth and computation power. Therefore it's essential to monitor the effectiveness of the caching system, then achieve a higher cache hit ratio accordingly. In short, the hit ratio determines how effectively the cache can reduce the overall memory-access time. The following figure illustrates the cache hit ratio (y-axis) estimation as a function of time (x-axis) on a daily basis.



Conclusion

TCPWave's [monitoring solution](#) helps our customers optimize infrastructure costs and effectively maximize speed with minimum latency. For a quick demo, contact the [TCPWave Sales Team](#).