ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

Understanding Network Latency in Thin Client Environment

Sumita U.Sharma[#1], Y.B.Gandole[*2]
[#]Department of Electronics, DCPE, HVPM. Amravati (India)
[*]Adarsha Science J.B.Arts and Birla Commerce, Mahavidyalaya, Dhamangaon Rly-444709 (India)

Abstract— ***The interest in the thin clients is increasing day by day, in the modern age, mainly because of frustration with the growing total cost of ownership of personal computers. The Thin Client also consumes very less power as compared to a regular personal computer and hence supports the Green Computing. But unfortunately, thin clients may not meet the usability goal of highly interactive response. For intensely interactive applications, the tight control of end-to-end network latency required by thin clients may be hard to guarantee at large scale. This is particularly the case with video and graphics-heavy applications and poses a challenge to the adoption of web-based services by organizations running global operations. Network latency causes a delay in transmitting a message from one location to another. This can be attributed to several other factors, such as network congestion, network traffic, and computer storage capacities. Of course, the distance between two locations is the main factor that contributes to the delay. Since transmission between two cities will not be a straight path, latency is subject to detour and can be a factor of any deviation between these cities. The proposed paper depicts the issue of Network Latency and its various aspects in a thin client network.***

Index Terms **— Bandwidth, Cloud computing, Network Latency, Thin Client.**

I. INTRODUCTION

Internet technologies have undergone radical developments and massive changes since its inception in 1960's. In the beginning, there were large challenges related to bandwidth and capacity. This led to research being focused on bandwidth sharing and congestion avoidance. In the last couple of decades, we have seen tremendous developments in networking technology, resulting in much higher bandwidths. This development is accompanied by a tendency among Internet users in general to consume much more bandwidth, both for uploading and downloading. The increase in bandwidth consumption is accelerated by peer-to-peer technology like Bit Torrent. Parallel to the trend of increased bandwidth usage on the Internet, real-time communication applications have also evolved and gained ground. Presently, the Internet is used as medium for a wide range of interactive services like chat, remote desktop, stock trade systems, IP telephony, video conferencing and networked games. The element of interactivity, however, leads to latency requirements; users become dissatisfied when they must wait for the system to respond. This is often problematic as the basic architecture of internetworking is based on best-effort services. Research has been performed, looking into ways of assuring a fixed quality of service (QoS) for data transport, but as the most successful approaches need support along the path of the connection, such approaches have not yet gained ground. A consequence of the lack of QoS mechanisms is that we still have to rely on end-to-end approaches in order to provide data.

Control techniques support low latency display operations on thin client devices. In response to a request to render a page in a user interface, the control may distinguish fields that likely have immediately displayable content from fields that are unlikely to have immediately displayable content. The control may retrieve data for those fields that are likely to have immediately displayable content and render them in an initial page. Content for the fields that are unlikely to have immediately displayable content may be generated as a background process and may be rendered in supplemental page(s) as they become available. The control permits quick rendering of basic pages, which may include navigation functionality, and therefore promote early navigation operations or quick review of data which can be made immediately available.

II. FACTORS THAT GENERALLY AFFECT THE LATENCY

Network bandwidth is just one element of what one can perceive as the speed of a network or internet. Latency is another element that contributes to network speed. The term latency refers to any of several kinds of delays typically incurred in processing of network data. A so-called low latency network connection is one that generally experiences small delay times, while a high latency connection generally suffers from long delays [1]. Though the maximum possible bandwidth of a network connection is theoretically fixed, according to the technology used, the actual bandwidth varies time to time and is affected by high latencies. Excessive latency

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013
creates bottlenecks that prevent data from filling the network pipe, thus decreasing effective bandwidth. The impact of latency on network bandwidth can be temporary and lasting a few seconds or persistent depending on the source of the delays. Besides propagation delays, latency also may also involve transmission delays that is the properties of the physical medium and processing delays like passing through proxy servers. In general, the factors that affect the network latency can be as follows...

A. Thin Client Hardware

A thin client is any diskless workstation with minimum hardware resources and it does not have hard drives or CD/DVD drives. A thin client has a less system memory or RAM as compared to a regular computer (thick client). The thin client totally relies on the server that provides all processing power, memory, data storage and all sorts of application software that the thin client uses. In other words a thin client is totally dependant on the server. Although a thin client works on bare minimum hardware resources, it does have standard specifications to yield the desired performance. The hardware requirements for the thin client are with respect to the processor it uses with its speed and the memory. The response time from the server for the request sent by client is mainly dependant on these factors particularly, if the response is in the form of high quality graphics and multimedia content, since the thin client needs to render the graphics and animations. This can be experienced in online or Internet gaming.

B. Internet Service Provider

Although the bandwidth doesn't directly affect the latency, the number of routers between client and the server can affect your ping times dramatically. Each router can introduce a routing delay so that a network performs better if it routes with the fewest number of hops.

C. Maximum Transmission Unit

Another often over-looked aspect is MTU or Maximum Transmission Unit. The MTU represents the largest single packet that can move from your network, through the ISP, to the destination. Generally it is around 1500, and most routers are set to this. However, if you're using DSL or some other technology with a packet overhead, you may need to lower your MTU to keep from having packet fragmentation. Packet fragmentation has a dramatic effect on the network speed. As a simple example, if the PC has defaulted to 1500, but the DSL line can only take 1492 (because it's 802.3), every single packet sent will be fragmented into one 1492-byte packet, and 1 8-byte packet. In such situation, the PC would send twice as many as packets wherein network collisions will be higher and also overhead will be higher, and the connection will suffer [2].

III. MEASURING NETWORK LATENCY

In almost all TCP-IP and Internet based applications like Mail, Web site, ftp, custom sockets applications in a network, latency matters just as much, or more, than bandwidth. If bandwidth is the “speed of our road”, then latency is the “journey time”, which is related to “the length of the road” as well as all the “hold-ups” (competing traffic, delay at routers, switches, repeaters etc) along the way [3]. If we want to test how an application will work in a real network using a network emulator then we need to understand the latency, bandwidth that it will experience.

Network tools like ping tests and trace route measure latency by determining the time it takes a given network packet to travel from source to destination and back, the so-called round-trip time or RTT. Round-trip time is not the only way to specify latency, but it is the most common. On DSL or cable Internet connections, latencies of less than 100 milliseconds (ms) are typical and less than 25 ms desired. Satellite Internet connections, on the other hand, average 500 ms or higher latency.

Excessive network latency creates bottlenecks in the network, which lead to a reduction of effective bandwidth. When determining network speed, it is thus important to use not only a bandwidth calculator, but also to run a network latency test.

A high variation of PING times on a data line (so-called “jitter”) can be a sign of an overloaded device or line. A well performing data line should show an almost steady value. Too many lost PINGs (“packet loss”) are also a sign of an overloaded network device [4].

A. Latency and traffic load

Latency can change as the traffic load changes. As load increases, it is possible that latency will increase since buffers may begin to populate on the path between the sender and receiver. Measuring latency while considering network load can get complicated. To fully characterize the latency versus load, measurements must be made at various network loads. Controlling the network load during the measurement is difficult fortunately; there are ways to simplify this process. Most enterprise networks have a fairly predictable bandwidth usage pattern. This pattern will change as new network applications are deployed and as more people use the network, but from one

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

day to the next there is little change in this pattern. Typically the network bandwidth used is low from late afternoon until the beginning of the business day. Network traffic will increase during the workday and there may be slight decline in utilization during lunch.

Knowing these patterns makes it possible to test network latency with known levels of network utilization. To systematically measure latency, it is good to sample latency throughout the day in regular intervals.

B. Latency Standards

Defining a level of latency that is deemed acceptable is difficult since it's hard to determine a threshold for user productivity based on application response times. However, it is important to define typical end-to-end latency values so you have a reasonable goal for latency. Monitoring is important since it is helpful to know typical latencies between each remote and the servers. An increase in end-to-end latency may be an indicator of a network problem. With these issues in mind, here are some "rules of thumb" for end-to-end latency (between workstation and servers):

- A round trip end-to-end latency of 30ms or less for LEAs is healthy. This measurement should be monitored to track any changes.
- Round trip latencies between 30ms and 50ms should be monitored. If there are ways to lower end-to-end latency, they should be considered.
- Round trip latencies greater than 50ms require immediate attention to determine the cause of the latency and possible remedies to lower the end-to-end latency. Monitor this measurement to track the improvements.

IV. SOFTWARE TOOLS FOR MEASURING NETWORK LATENCY

SmokePing is a valuable tool for Linux/UNIX and is similar to the Multi Router Traffic Grapher (MRTG). SmokePing measures the latency and packet loss of your network and also has alarm features that can be configured to provide real-time notifications, as well as historical data. You can use it to monitor WAN interfaces or LAN destinations. SmokePing is an excellent freeware product written by Tobi Oetiker. It relies on his powerful RRD TOOL, which is a Round Robin Database back end that provides easy-to-use, fast, compact, time-based data collection and display capabilities. The output of SmokePing is an easily modifiable .html file.

Smokeping is a program that looks the latency of the connection between your servers and multiple remote hosts. SmokePing is a deluxe latency measurement tool. It can measure, store and display latency, latency distribution and packet loss. SmokePing uses RRDtool to maintain a long-term data-store and to draw pretty graphs, giving up to the minute information on the state of each network connection. Fig. 1 shows a standard user interface and the output of the SmokePing software.

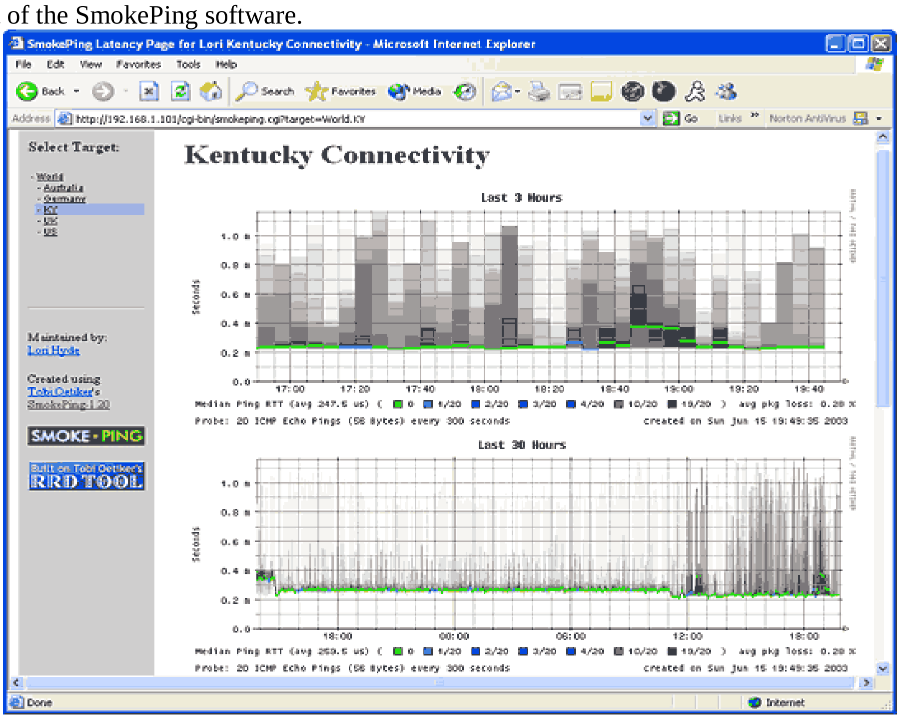

FIG. 1: SMOKE TEST USER INTERFACE SCREEN

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

V. CONCLUSION

Issues with latency and dropped packets can kill a network's performance and cripple applications like real-time communications, scientific computing, online gaming, video streaming and high-frequency trading. But the problems can be extremely difficult to diagnose, as they may not appear under test conditions, and real-time monitoring of performance can require dedicated hardware or procedures that actually cut into the usable bandwidth. A team of academic researchers have come up with what they think is a solution, one that could sample the transmission of a collection of representative packets in real time, in a manner that's inexpensive in terms of both hardware and networking resources.

REFERENCES

[1] http://superuser.com/questions/129386/factors-affecting-network-latency.

[2] http://blogs.itrinegy.com/2009/02/how-do-you-measure-latency-rtt-in-a-network-these-days/.

[3] http://www.paessler.com/network_latency.

[4] Quantifying Interactive User Experience on Thin Clients, Niraj Tolia, David G. Andersen, and M. Satyanarayanan Carnegie Mellon University.

[5]] http://smutz.us/techtips/NetworkLatency.html.

[6] http:// compnetworking.about.com/od/speedtests/a/ network_latency.htm.