Research on the Distribution and Self-Similarity Characteristic of End-To-End Network Delay

Min Liu¹, Yanru Xue¹, Yang Zhao²* and Huiling Guo³

¹Hebei Normal University of Science & Technology, Qinhuangdao 066004, China
²Department of Electronic and Information Technology, Jiangmen Polytechnic, Jiangmen 529090, China
³College of Computer Science and Technology, Zhoukou Normal University, Henan 466001, China
zhaoyang_2015@sina.com

Abstract

As an important indicator of evaluating network performance, network delay can reflect the transmission performance of the current path and also the service level provided by the opposite terminal host. In the paper, analysis is first conducted on the composition of network delay; besides, through Ping measuring method, Ping measuring network delay experiment is conducted on different destination hosts of local area network (LAN) and wide area network (WAN). Besides, the distribution features of network delay as time changes is analyzed, while the influence of data package (in terms of the size) on network delay will be explored in the paper. After the process, the definition, along with the distinguishing method of the self-similarity process is introduced, and self-similarity distinguishing is rendered on network delay through the variance-time plot. According to analysis result, network delay is featured by strong nonlinearity and self-similarity. Compared with LAN, WAN is endowed with higher long-range dependency.

Keywords: S-Ping, Network time delay, RTT, Measurement, Self-similarity

1. Introduction

Parameters for evaluating network transmission performance include network connectivity, transmission delay, throughput, network bandwidth and the like. Among them, network delay is a relatively important evaluation parameter. Network delay is affected by multiple factors such as network topology, number of forwarding/routing nodes, route processing performance, routing algorithms and background traffic, which reflects the transmission performance of current path and the level of service provided by opposite end host, and changes randomly over time [1]. Through network delay, network load conditions of current link and the degree of network communication congestion can be reflected.

As one of the important indexes of network performance of network time delay, research is of great significance, the network communication system, network control system, remote medical treatment, network video transmission has a direct impact [2-4]. Besides the Ping method [5, 6], there are some other methods which can be adopted for network delay measurement. In 2006, facing the clock synchronization during the network measurement process, Sun Haiyan et al., [7] put forward with the optimization method based on the minimum normal direction, which can be utilized to estimate the clock error between the two points, thus obtaining a more real measuring result. In 2008, through the server-client connection mode, Zhong Zhiyan [8] measured the TCP
transmission delay between the two points of the LAN with the assistance of QCheck network testing tool. Through the method, the characteristics of the time delay between the server and client can be precisely reflected. However, network transmission of LAN can hardly be affected by external disturbance, due to which the data acquired could be unduly ideal, and in lack of universality. Wu Pingdong research group [9], from Beijing Institute of Technology, has conducted a deep analysis on the network delay features and the causes during the process of remote control, pointing out that network delay distribution is provided with certain regularity.

As we know, network provides is the Best Effort data transmission service. Network structure is in a dynamic and changing state, and network flow is featured by disequilibrium. Hence, network delay is uncertain and changing dynamically. Besides, network performance measurement serves as an important precondition for accuracy control concerning the network. Meanwhile, with the existence of network congestion control mechanism, the route queuing time and network load will exert great influence on point-to-point time delay. Based on the initial network delay modeling, time delay sequence is featured by short correlation property, and the modeling for network time delay can be realized through Markov model [10] or regression model [11]. However, the latest research results in recent years show that, the distribution of network time delay presents a feature of self-similarity and long-range correlation[12-14]. Through the analysis on Internet network time delay data, Borella[13] holds the idea that time delay is provided with long-range correlation.

In the literatures, the measurement of network time delay is conducted through Pchar tool, while the test results have verified the result that network time delay is in accordance with Pareto distribution pattern [15], namely the feature of self-similarity. Based on the research, Zhou Xiaobing found that distribution with self-similarity can better describe network time delay; that is, network time delay can be provided with statistics similarity in all time scales. In terms of time series with short correlation property, with the increase of time interval, self-correlation function poses an exponential decay and correlation decreases rapidly; however, in terms of time series with long correlation property, with the increase of time interval, self-correlation function declines in a hyperbolic function with comparatively slower decaying rate. Hence, in order to conduct such analyses as time delay modeling and prediction on network time delay, the self-similarity feature of network time delay shall be analyzed at first. In the paper, time delay measurement is conducted from the perspective of real physical link within LAN and WAN, thereby analyzing the self-similarity of network delay in both the idle state and busy state of the network.

The structure arrangement of this paper is as follows: in the following section, the definition and major components of network delay are introduced while in the third section, through the Ping measuring method, time delay data of LAN and WAN are obtained and the distribution character of time delay data is analyzed. In the fourth section, the definition, along with the distinguishing method of self-similarity is introduced and self-similarity distinguishing is rendered on network delay based on the variance-time plot.

2. End-to-end Delay Compositions

Generally speaking, end-to-end network delay consists of four parts: propagation delay, transmission delay, queuing delay and processing delay. Among them, transmission delay, propagation delay and processing delay are deterministic delays, whereas queuing delay is a random delay. Therefore, network delay is overall a random variable. Figure 1 is a diagram showing the composition of network delays between three routing nodes in transport layer channel.
1) $TD_{prop}$ (Propagation Delay), segment AB in Figure 1 refers to the time taken from sending of the first bit of data packet by data acquisition unit to receipt of the bit by server. The propagation velocity of electromagnetic wave is about $2/3$ the speed of light in network transmission cable, hence the propagation delay can be described by Formula (1):

$$TD_{prop} = \frac{d}{v}$$  \hspace{1cm} (1)

Where $d$ denotes the communication distance (unit: m) between data acquisition unit and server, and $v$ represents the propagation velocity of electromagnetic wave in the communication link (unit: m/s).

2) $TD_{tran}$ (Transmission Delay), s, segment BC in Figure 1 refers to the time taken from entrance of the first bit of data packet into the communication link to entrance of the last bit of the data packet into the communication link. The calculation process of transmission delay is described in Formula (2):

$$TD_{tran} = \frac{L}{C}$$  \hspace{1cm} (2)

Where $L$ denotes the size of data packet (unit: byte (1byte=8bit)), and $C$ represents the bandwidth of network (unit: bit/s).

3) $TD_{queu}$ (Queuing Delay), s, segment CD in Figure 1 refers to the time taken for the data packet to queue in the router buffer, which is mainly affected by router performance and network load. In general, the value of queuing delay changes randomly.

4) $TD_{proc}$ (Processing Delay), s, segment DE in Figure 1 refers to the time taken from arrival of data packet at input end of a certain node in the network to departure of that node, which is mainly affected by the computing power of that processing node.

3. S-Ping Method

Conventional network Ping measurement approach could only obtain four sets of data per measurement, which can hardly record measurement time point; besides that, the time interval of Ping is fixed as 1s. To be able to record in detail the time point information on measurement delay of each set of Ping, and measure the network delay from multi-cycles, we develop an improved Smart Ping (S-Ping) measurement method based on Visual Basic
4. Script (VBS) scripting language. The S-Ping allows recording of detailed measurement time nodes, and changes of size of probe packet and probing cycle. A complete S-Ping record is as shown below:

Microsoft (R) Windows Script Host Version 5.6
(C) Microsoft Corporation 1996-2001

2013-5-8 20:47:00 Reply from 121.194.0.239: bytes=64 time=50ms TTL=52
2013-5-8 20:48:00 Reply from 121.194.0.239: bytes=64 time=56ms TTL=52
2013-5-8 20:49:00 Reply from 121.194.0.239: bytes=64 time=54ms TTL=52
2013-5-8 20:50:00 Reply from 121.194.0.239: bytes=64 time=60ms TTL=52

4. S-Ping based Network Time Delay Measurement Experiments

To observe the distribution characteristics of network delay on different time scales, and study the impact of probe packets of different sizes on network delay, we design a total of three sets of S-Ping delay measurement experiments, which will be presented in detail.

Here, we do S-Ping measurements from a computer in Room 425 of Civil Engineering Building Four in Zhoukou Normal University (ZNU) to four different target hosts. They are host in library of ZNU, host in data center in Zhengzhou, host of Sina in Beijing, host of Google in American, respectively. The further information of these destination hosts is shown in Table 1. As can be seen, the types of network include LAN and WAN, communication distance ranges between 1km~15,000km, and hop of intermediate routing nodes is between 4~20.

<table>
<thead>
<tr>
<th>Target host</th>
<th>Library</th>
<th>Data center</th>
<th>Sina</th>
<th>Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP address</td>
<td>211.67.150.2</td>
<td>202.196.64.199</td>
<td>121.194.0.239</td>
<td>74.125.128.199</td>
</tr>
<tr>
<td>Number of routers</td>
<td>4</td>
<td>20</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td>Beijing</td>
<td>USA</td>
</tr>
<tr>
<td>Estimated distance/km</td>
<td>1</td>
<td>400</td>
<td>1000</td>
<td>15000</td>
</tr>
</tbody>
</table>

4.1. Periodic Property of Time Delay

Experiment purpose: Measuring the periodical characteristics of network time delay in long time scales.

Experiment period: from midnight, 15/5, 2014 to midnight, 22/5, 2014.

Target host: Data center in Zhengzhou and Sina server in Beijing.

S-Ping action: The probe packet size is 32B, the ping period is 15 minutes, and obtain 672 sets of data.

Figure 2 shows the measurement results. It can be seen that, network delay has an obvious periodicity in a week's time scales. The value of network delay is larger and fluctuates relatively violently during daytime, while during night and early morning period, the value of network delay is lower and fluctuates relatively smoothly. It is worth noting that network delay exhibits no significant difference between weekday and weekends periods.
In Figure 2 (a), network delay of Ping data center is 433ms at maximum, and 49ms at minimum, with a mean of 148ms, however, in Figure 2 (b), network delay of Ping Sina has a maximum of 95ms, and a minimum of 17ms, with a mean of 36ms. Although the data center, which is located in Zhengzhou, is closer to the test hosts in Zhoukou compared to Beijing's Sina server, its average delay is much larger than Sina.

The cause of this phenomenon is that despite the closer distance between Zhengzhou and Zhoukou, the number of routing nodes between them is 20, which has 8 more nodes compared with the Sina servers located in Beijing, so the delay of queuing in routing nodes during transmission of data packet is greater than the propagation delay. In addition, network nodes Sina server is located are mostly backbone nodes, whereas the Zhengzhou data center servers are just ordinary nodes in the network; in comparison, routing performance and processing speed of backbone network nodes are faster.

![Figure 2. Continuous Ping Experiments](image)

**Figure 2. Continuous Ping Experiments**

4.2. Time-variant Characteristic of Time Delay

*Experiment purpose:* Measuring the distribution of network time delay over time.

*Experiment period:* from midnight, 23/5, 2014 to midnight, 24/5, 2014.

*Target host:* All the four target hosts.

*S-Ping action:* The probe packet size is 32B, the ping period is 1 minutes, and obtain 1,440 sets of data.

The statistical features of network time delay of data as shown in Table 2. From the perspective of values of delay, the minimum is happened in pinging library (LAN), with a mean of only 1.53ms, and Google server, which is farthest away, has the largest delay, with a mean of 372.07ms. Although the communication distance of Sina server, which is located in Beijing, is farther than the data center of data center in Zhengzhou, its mean delay is only 40.09ms, which is much smaller than 155.84ms, the mean delay of the data center. This conclusion in conformity with the previous section. On the other hand, from the point of view of packet loss, it can be clearly seen that packet loss in pinging data center is rather severe, reaching 7.78%; in comparison, Google server has a packet loss rate of only 1.11% despite its farther distance and mean delay of 372.07ms.
Table 2. Statistics of Network Delay Measurement

<table>
<thead>
<tr>
<th>Target host</th>
<th>Library</th>
<th>Data center</th>
<th>Sina</th>
<th>Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of packet loss</td>
<td>15</td>
<td>112</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>Packet loss rate</td>
<td>1.04</td>
<td>7.78</td>
<td>0.42</td>
<td>1.11</td>
</tr>
<tr>
<td>Min time delay /ms</td>
<td>1</td>
<td>54</td>
<td>17</td>
<td>54</td>
</tr>
<tr>
<td>Max time delay /ms</td>
<td>16</td>
<td>539</td>
<td>84</td>
<td>882</td>
</tr>
<tr>
<td>Avg time delay /ms</td>
<td>1.53</td>
<td>155.84</td>
<td>40.09</td>
<td>372.07</td>
</tr>
</tbody>
</table>

The main reasons for above situation are as follows:

LAN-to-LAN communication features short distance and stable network, and therefore has minimal network delay and less packet loss.

Sina and Google servers are both distributed farther, but as they are in the backbone network nodes, their network communication link is relatively stable, so their packet loss rate is relative low.

Compared to Sina and Google, data center has a closer distance to the client side, but the number of routing nodes between them reaches as many as 20, which increases the corresponding queuing and processing delays, so when pinging data center, there occurs larger delay, and more packet loss.

Figure 3 (a)–(d) show the curves of measured pinging network delays between client and the four destination hosts, respectively. As can be seen, network delays are distributed randomly within a certain range, for the library which is within LAN, network delays are mostly located in the vicinity of the mean; for Zhengzhou data center and Sina server, delay presents a periodicity over time, within daytime, network load is larger, delay value is higher, and delay fluctuation is greater, while during late night and early morning, delay value is lower, and delay fluctuation is relatively smooth; for Google server, due to the time difference between China and the United States, network load changes constantly over time, so delay jitter is more obvious, and fluctuation range is relatively large.

![Figure 3. The Network Delay Measurement Curve]
4.3. The Effect of Probe Packet Sizes on the Network Delay

Literature [16] believes that end-to-end network delay is linearly related to the size of probe packet; the bigger the size of probe packet, the greater the value of network delay. To verify the impact of probe packet size on network delay, Experiment 3 is performed as shown below.

**Experiment purpose:** To study the effect of detecting packet size on the network delay.

**Experiment period:** 18:00, 28/5, 2014 to 20:00, 28/5, 2014.

**Target host:** All the four target hosts.

**S-Ping action:** The probe packet size is 32B, 64B, 128B, 256B, 512B and 1,024B, respectively, the ping period is 10 seconds, and obtain 1,440 sets of data in six different probing packets type.

Table 3 describes the statistical features of this experiment. As can be seen from the table, within 32B~1024B, the maximum value of average delay tends to occur at 1024B (512B for Google), but no particularly significant differences are noted. It is also similar for packet loss rate; packet loss rate for a 32B data packet shows little difference from a data packet with a size of 1,024B. The actual measured results of this section are contrary to the simulation results reported in literature, the reasons may be attributable to the network model in literature, and the impact of background traffic. It can be seen from the analysis of network delay composition that the bigger the data packet, the larger the transmission delay, but when data packet is smaller than 1460B, i.e., maximum transmission unit (MTU) of TCP, its impact on delay is not obvious.

**Table 3. Statistics of Network Delay Measurement with Different Probing Packets**

<table>
<thead>
<tr>
<th>Target host</th>
<th>performance</th>
<th>32B</th>
<th>64B</th>
<th>128B</th>
<th>256B</th>
<th>512B</th>
<th>1024B</th>
</tr>
</thead>
<tbody>
<tr>
<td>library</td>
<td>Number of packet loss</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Packet loss rate</td>
<td>0.07</td>
<td>0.07</td>
<td>0.28</td>
<td>0.21</td>
<td>0</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Min time delay /ms</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Max time delay /ms</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Avg time delay /ms</td>
<td>1.00</td>
<td>1.01</td>
<td>1.00</td>
<td>1.02</td>
<td>1.02</td>
<td>1.04</td>
</tr>
<tr>
<td>Data center</td>
<td>Number of packet loss</td>
<td>277</td>
<td>269</td>
<td>255</td>
<td>277</td>
<td>243</td>
<td>297</td>
</tr>
<tr>
<td></td>
<td>Packet loss rate</td>
<td>19.2</td>
<td>18.7</td>
<td>17.7</td>
<td>19.2</td>
<td>16.9</td>
<td>20.6</td>
</tr>
<tr>
<td></td>
<td>Min time delay /ms</td>
<td>140</td>
<td>141</td>
<td>151</td>
<td>150</td>
<td>140</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>Max time delay /ms</td>
<td>205</td>
<td>209</td>
<td>212</td>
<td>206</td>
<td>210</td>
<td>210</td>
</tr>
<tr>
<td></td>
<td>Avg time delay /ms</td>
<td>155.79</td>
<td>155.80</td>
<td>156.00</td>
<td>155.95</td>
<td>156.51</td>
<td>156.98</td>
</tr>
<tr>
<td>Sina</td>
<td>Number of packet loss</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Packet loss rate</td>
<td>0.07</td>
<td>0.14</td>
<td>0</td>
<td>0.28</td>
<td>0.21</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Min time delay /ms</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Max time delay /ms</td>
<td>23</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Avg time delay /ms</td>
<td>15.34</td>
<td>15.37</td>
<td>15.38</td>
<td>15.56</td>
<td>15.95</td>
<td>16.38</td>
</tr>
<tr>
<td>Google</td>
<td>Number of packet loss</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Packet loss rate</td>
<td>0</td>
<td>0.14</td>
<td>0.07</td>
<td>0.28</td>
<td>0.28</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Min time delay /ms</td>
<td>62</td>
<td>59</td>
<td>63</td>
<td>62</td>
<td>65</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Max time delay /ms</td>
<td>347</td>
<td>344</td>
<td>350</td>
<td>342</td>
<td>347</td>
<td>347</td>
</tr>
<tr>
<td></td>
<td>Avg time delay /ms</td>
<td>179.22</td>
<td>178.49</td>
<td>180.40</td>
<td>180.85</td>
<td>181.32</td>
<td>180.63</td>
</tr>
</tbody>
</table>

Figures 4~7 show the curves of measured network delays as well as smooth curves of delays when probe packets of different sizes ping the four destination hosts, respectively. As can be seen from the delay measurement curves on the left side, delay has a random distribution feature; difference in delay is rather small among probe packets of different sizes.
sizes, delays tend to fluctuate within a certain range, accompanied by occasional sharp jitter. It can also be seen from the delay smooth curves on the right side that the size of probe packet does not have a great impact on delay, smooth curve trend is obvious only for pinging Sina server, which also displays increase in delay value with increasing probe packet size. But the data of remaining three groups does not present this feature, on the contrary, partially enlarged views of Figures 4 (b) and 7 (b) show maximum delay value at 512B.

**Figure 4. Network Delay and its Smooth Curves of Different Packets Ping Library**

**Figure 5. Network Delay and its Smooth Curves of Different Packets Ping Datacenter**

**Figure 6. Network Delay and its Smooth Curves of Different Packets Ping Sina**
5. Self-similarity Characteristics of Network Delay

5.1. Definition of Self-similarity

Let $X = (X_t : t = 0, 1, 2, \ldots)$ be a covariance stationary random process, its mean value $\mu$ be constant, its variance $\sigma^2$ be finite variance and autocorrelation function be $r(k), \ k \geq 0$. At the same time, $r(k)$ is only correlated with $k$.

$$r(k) = \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{\sigma^2} \quad (k = 0, 1, 2, \ldots)$$

Among this, supposing that the autocorrelation function $r(k)$ satisfies

$$r(k) \propto k^{-\beta}L(t) \quad k \to \infty$$

where $0 < \beta < 1$ and $L(t)$ is a function slowly varying with time

$$\lim_{t \to \infty} \frac{L(tx)}{L(t)} = 1 \quad x > 0.$$  

Let $X^{(m)} = \{X^{(m)}_k : k = 1, 2, 3, \ldots \quad m = 1, 2, 3, \ldots \}$ be another covariance stationary time sequence, among which $X^{(m)}_k$ can be figured out through formula (5).

$$X^{(m)}_k = \frac{X_{km-m+1} + \cdots + X_{km}}{m}$$

**Definition 1**[17]: Process $X$ is called a strict second order self-similarity process and the self-similarity parameter is $H = 1 - \frac{\beta}{2}$, where $0 < \beta < 1$, if following two conditions are satisfied at the same time as shown in (6).

$$\begin{align*}
\text{var}(X^{(m)}) &= \sigma^2 m^{-\beta} \\
r^{(m)}(k) &= r(k) \quad k \geq 0
\end{align*}$$
Definition 2 [17]: Process $X$ is called a gradual second order self-similarity process and the self-similarity parameter is $H = 1 - \frac{\beta}{2}$, where $0 < \beta < 1$, if all $k$ which is big enough satisfy (7).

$$r^{(m)}(k) \rightarrow r(k) \quad m \rightarrow \infty \quad (7)$$

$H$ (Hurst) in the definition is the self-similarity parameter.

5.2 Discriminant Methods of Self-similarity

As the self-similarity process has to satisfy $0.5 < H < 1$, the self-similarity of the time delay sequence can be described by the Hurst parameter. If the value of $H$ is larger, it means that the self-similarity of the network delay is higher and the burstiness of time delay sequence is stronger. Hurst parameter has many estimation methods, which can be grouped mainly into time domain analysis and frequency domain analysis. The details are shown in Table 4.

Table 4. Estimation Methods for Hurst

<table>
<thead>
<tr>
<th>Classification</th>
<th>Name</th>
<th>algorithm complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Domain Methods</td>
<td>Variance-Time</td>
<td>$O(N)$</td>
</tr>
<tr>
<td></td>
<td>R/S method</td>
<td>$O(N)$</td>
</tr>
<tr>
<td></td>
<td>Absolute method</td>
<td>$O(N^2)$</td>
</tr>
<tr>
<td></td>
<td>Residue method</td>
<td>$O(N^2)$</td>
</tr>
<tr>
<td>Frequency Domain Methods</td>
<td>Whittle estimation</td>
<td>$O(N^2)$</td>
</tr>
<tr>
<td></td>
<td>periodogram method</td>
<td>$O(N \log N)$</td>
</tr>
<tr>
<td></td>
<td>wavelet analysis method</td>
<td>$O(N \log N)$</td>
</tr>
</tbody>
</table>

In addition to above-mentioned methods, there are also many improved algorithms based on the primary algorithm. The main aim of this is to improve the accuracy of the estimation algorithm or the calculating speed of the algorithm. Among these methods, variance-time plot (V-T, Variance-Time) is a kind of clustering variance method. This method, of which the calculating speed is very fast and the time complexity is $O(N)$, is usually employed to determine whether the time sequence has self-similarity. Therefore, variance-time plot is adopted in this paper to verify the self-similarity of time delay sequence under different network conditions. It can be known from formula (6) that for a larger $m$, the variance of the self-similarity compressed time sequence can be expressed by formula (8):

$$Var(x^{(m)}) \sim \frac{Var(x)}{m^\beta} \quad (8)$$

Take the logarithm of both sides of formula (8), following can be acquired:

$$\log\left[Var(x^{(m)})\right] \sim \log[Var(x)] - \beta \log(m) \quad (9)$$
where, $\log[\text{Var}(x)]$ is a constant correlated with the value of $m$. If showing the point of $(m, \text{Var}(x^m))$ in the log-log plot, then a curve with the slope of $-\beta$ can be obtained and the self-similarity parameter can be worked out by $H = 1 - \frac{\beta}{2}$.

5.3. Self-similarity of the Time Delay Sequence

The (a) ~ (d) in Figure 8 correspond to the self-similarity parameter estimation diagrams of the network delay of four cases, library, data center, Sina and Google, respectively. It can be seen from the diagrams that self-similarity emerges in all measuring results of the four conditions and their values of Hurst parameter are $H_a=0.6984$, $H_b=0.7251$, $H_c=0.8511$ and $H_d=0.9386$, respectively. Thus, it can be known that the self-similarity of WAN is higher than that of LAN and meanwhile, the self-similarity of time delay in the busy state of the network is higher than that in the idle state of the network, showing that WAN is endowed with higher long-range dependency.

![Figure 8. Network Topology Connection Diagram of Four Cases](image)

6. Conclusion

End-to-end network delay mainly consists of four parts, namely transmission delay, propagation delay, queuing delay and processing delay, among which the queuing delay of routing node occupies the major part of delay composition. It can be seen through three sets of S-Ping network delay measurements that network delay has an overall periodic distribution feature, network delay is larger and fluctuates markedly during working
hours, while during late night and early morning hours, delay changes mildly within lower values. Besides, compared with the queuing delay, the impact of propagation delay caused by physical distance is very small. When the size of probe packet is within the MTU range, network delay does not have a significant linear relationship with probe packet size, which provides the basis for the selection of probe packet size for measurement of network delay between data acquisition unit and server.

The nonlinearity feature of network delay is verified through the measured data. In addition, the self-similarity of network delay is analyzed emphatically in this paper. According to the result, the self-similarity of WAN is higher than that of LAN and the self-similarity of time delay in the busy state of the network is higher than that in the idle state of the network, showing that the long-range dependency of WAN is higher. The study of this paper is helpful to the establishment of an appropriate network delay model, which can lay a solid foundation for the prediction of network delay.

References