Scalable Application-Specific Measurement Framework for High Performance Network Video

Congxiao Bao
Network Center of Tsinghua
University, Beijing, 100084, China
86-10-62603020

congxiao@cernet.edu.cn

Xing Li
Network Center of Tsinghua
University, Beijing, 100084, China
86-10-62785983

xing@cernet.edu.cn

Jinpeng Jiang
Dept. of Electronic Engineering of
Tsinghua University, Beijing, 100084,
China

Jjp02@mails.tsinghua.edu.cn

ABSTRACT

Extensive studies show that bandwidth provisioning is the key method to guarantee the performance of mission-critical real-time network applications. Unfortunately, network congestions, rate-limiting for special addresses and ports, and even the bad cabling may result in the bandwidth bottleneck along the end-to-end path. This is the main cause of degradation of the application performance, especially for the high-performance video applications which consumes large-bandwidth. Our experiences, being both as a national academic Internet Service Provider and end-user of high-performance video applications, clearly indicate that using suitable measurement tools are critical to find the network bottlenecks and enable a successful video session.

To support high-performance video on the Internet, taking the popular Digital Video Transport System (DVTS) as an example, we propose a Scalable Application-Specific Measurement (SASM) framework. The SASM for DVTS involves two novel components. One is a set of DVTS specific network and application measurement tools, which provide a common understanding for both end-users and network administrators. The other is the scalable management scheme for distributed DVTS-specific measurement servers, including server auto-discovery and server auto-registration. Deploying SASM on several academic backbones indicates the effectiveness of the proposed framework for the troubleshooting and network fine-tune of high-performance video applications.

Categories and Subject Descriptors

C.2.3 [Computer-Communication networks]: Network Operations

General Terms

Management, Measurement, Performance, Design, Experimentation

Keywords

High performance network video, application-specific measurement, measurement server auto-registration, measurement server auto-discovery

1. INTRODUCTION

More and more demonstrations and trials of real-time high-performance video applications on the Internet have been carried on for recent years. Digital Video Transport System (DVTS) [16] is the most attractive and useful one among those systems. With relative high-resolution (720*480 pixels), reasonable latency and acceptable setup cost, DVTS is widely used for global distance-learning on telemedicine [10], music-teaching and art performance collaboration among academic communities in Asia-Pacific, Europe and North America. However, due to the requirement of 30Mbps for a single DVTS stream, setting up DVTS system even for a point-to-point connection is a big challenge for both Internet Service Providers (ISP) and end-users, especially for inter autonomous system (inter-AS) collaboration.

Besides the performance problems described in [2], there are several factors which make the 30Mbps transmission difficult on the Internet.

(1) The link bandwidth is limited and the link utilization is high in some segments of the packet traveling path.
(2) The network administrators may setup firewalls and QoS features to suppress the large throughput for specific protocols, IP addresses, port numbers, packet size and even the application headers for security reasons.
(3) Ethernet devices may have bad cabling and/or wrong duplex mode setup.
It is clear that a single breakpoint on the video traveling path will cause degradation of application performance or even destroy the application session. Furthermore, network behavior may change between two events or even two rehearsals. It is possible for the network administrator to configure the BGP routing to guarantee the bandwidth provisioning for one specific case, but it may not be suitable for others. Doing routing adjustment case by case also increases the possibility of wrong configuration by the administrators when a lot of requests arrive in short period.

In addition, setting-up high-performance video applications on Internet like DVTS is a complex process which involves network protocols, application protocols and human protocols working together [1]. The end-to-end video performance is often plagued by application, operating system and network problems interacting together. Heavy load of negotiation among end-users, application system engineers and network administrators, time differences and sleepless working for all-sites are the non-technique factors which make the test painful.

Existing measurement tools for troubleshooting the high performance video applications are mainly divided into two categories, the network performance measure tools and the

application-related trouble-shooting tools. Usually, the network performance measurement tools, like ping, traceroute [8], pathping [20], AMP/iperf [9] [11][17] etc. measures network performance using different protocol, port numbers and traffic patterns from the real video applications. Therefore, the measurement statistics based on these tools may not reflect the real network and application evidence. H.323 Beacon [2] [3] is a pioneer of application-related troubleshooting tool. It adds additional features sufficed in the context of H.323 application, however, the bandwidth requirement of H.323 video application is less than 3 Mbps. Our experiences also indicate that the main problems of H.323 video application occurred in the "last mile", (i.e. the firewall and end-system configuration), while the main problems of DVTS application occurred in inter-AS borders under complicated routing policies. Furthermore, most of the existing tools lack the ability for a scalable management scheme for large-scale distributed measurement. The studies in [12][13] outline the large-scale distributed information management scheme based on a mathematic model, but those descriptions only give general guidelines for the design and implementation of distributed contents without specific technologies on distributed servers management in detail. To meet the high-performance video application requirements, we do need suitable measurement tools. The design goals of such tools include:

(1) The measurement tools should be able to simulate the real video application characters, i.e. the protocol type, the source and destination port number and the traffic patterns, in order to provide the real network and application evidence for troubleshooting and isolate the application problems from network problems.

(2) Since the hop by hop performance guarantee is the basis for end-to-end performance, the measurement server should be placed alone the packet traveling path. So, the end-to-end performance problems can be decomposing into segment or even hop problems. The measurement tools should be able to find these servers and to do the measurement test with each server automatically at any time.

In this paper, taking DVTS as an example, we propose a new measurement framework supporting large-throughput real-time video applications, called the Scalable Application-Specific Measurement framework (referred to as SASM in the rest of the paper). In SASM for DVTS, we developed a set of DVTS-specific network and application measurement tools (dvping, dvmcast and the existing DVTS application), which can provide a common understanding among end-users, application system engineers and network administrators. A simple but straightforward server management scheme, including measurement server auto-registration, intra-autonomous system measurement and measurement server auto-discovery is also proposed in this paper.

The rest of the paper is organized as follows. Section 2 describes the key components and general framework of SASM. Section 3 introduces the case study of SASM in CERNET [14], CERNET2 [15] and TEIN2 [18]. Finally, concluding remarks and future work are given in Section 4.

2. SASM FRAMEWORK

2.1 Application-Specific Measurement Tool

The SASM uses general client/server architecture. The testing between two segments of the network can be achieved by enabling a communication session of a pair of client-server in the corresponding segments. It is important that the communication socket of a client-server pair for measurement tool must have the same socket parameters (protocol, destination address, destination port, source address, source port, application payload pattern) as that of a real DVTS streaming. As shown in Figure 1, the server called dvmcast has two modes (sending and reflecting) and the DVTS application and dvping are the two types of clients. The dvmcast can either read a DVTS data file [5] [6] and send it to the client, or reflects the received real-time DVTS streaming from a client and send back to the same client. The DVTS application system, which receives the video stream by dvmcast or sends a video stream to dvmcast and receives the reflected video stream, can be used by the end-users to check the video/audio quality subjectively. DVTS file sending on demand also provide a possible way for the network administrators to use and understand what the application users are talking about. The dvping, which sends a simulated DVTS stream with the same UDP protocol, destination address, destination port of DVTS session, source address, source port of DVTS session, DVTS payload pattern [5] to dvmcast and receives the reflected packet stream with packet loss, round trip time, sending and receiving throughputs reporting, can be used by both the end-users and the network administrators to check the network performance. Based on this network and application measurement combination, the end-users and the network administrators can exchange the testing results without the help of the counterpart and understand the problems described either in the network engineering language in terms of loss/ throughput/latency or the application language in terms of video/audio quality. Since intentionally or unintentionally sending DVTS stream from dvmcast can generates large traffic congestion, our systems are designed that the streaming can only be send to a permitted hosts. For the same reason, dvping has a special feature, which automatically limits the sending throughput in the bad network condition. Therefore, the DoS attack can be avoided and the impact of the proposed measurement testing to network can be minimized.
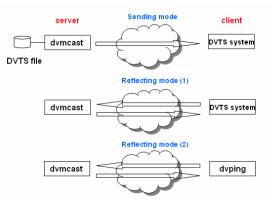

Figure 1 DVTS-specific measurement tool

2.2 Scalable Management Scheme for Distributed Servers

To guarantee the end-to-end video application performance, it is necessary to do troubleshooting segment by segment or hop by hop using distributed servers.

2.2.1 Methodology for the scalability

Currently there are about 1,342,177,280 addresses in the global Border Gateway Protocol (BGP) routing table [4]. This is roughly the theoretical maximum number of hosts on the Internet, which may run end-to-end high-performance video applications. Therefore, we do need a method to handle the scalability problem for the measurement test. It is well-known that the scalability problem can be solved by aggregation. Figure 2 shows the different numbers of possible measurement servers (referred to as MS in the rest of paper) required based on the different aggregations methods on the Internet. The tradeoff is the better aggregation (scalability) or the better measurement accuracy along the path. It is usually true that the network performance bottleneck happens at the autonomous boundary or the access networks, therefore, we chose to pickup at least one measurement server in each autonomous system. Based on this method, the number of measurement servers should be reduced to 20,000, the same as the current number of autonomous systems. Any IP address on the Internet can be mapped to the autonomous system number (referred as ASN in the rest of paper) using the global BGP table.

aggregation method	number of MS required	better aggregation	better measurement accuracy alone the path
a single MS	1		
MS in ASN confederation	2,000+		
MS in each ASN	20,000+		
MS in each prefix in BGP	200,000+		
MS in each subnet	2,000,000+		
MS in each host	100,000,000+		

Figure 2 Estimation of different level of aggregation and possible servers

2.2.2 Intra-AS measurement

In order to follow the rule of "at least one MS in each AS is needed", the assumption "the network performance is reasonably good inside an autonomous system" have to be ensured. We suggest deploying measurement servers on every possible Pop inside the same autonomous system and propose to run dvmcast and dvping inside each autonomous system periodically. The visualized performance reports of full-mesh tests among Pops with packet loss and latency based on beacon method [7] is called dvbeacon in SASM. The network administrators should be able to monitor the dvbeacon and fine-tune the network in their autonomous system to guarantee the above assumption. Figure 3 gives an example of performance monitoring of dvbeacon inside one autonomous system.
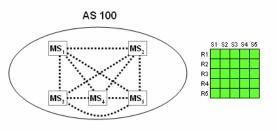

Figure 3 Performance monitoring inside an AS

2.2.3 Mapping

Mapping is the process which involves two sub-processes, one is any IP address of each hop obtained by traceroute to the ASN mapping and the other is ASN to measurement server IP address mapping in the same autonomous system. The explanation of two mapping procedures is shown in Figure 4.

The first mapping is based on the global BGP routing table, either obtained from a local BGP speaking router or obtained from global BGP routing table lookup services (e.g. routeviews provided by the University of Oregon's Route Views Project [19]). For example, in order to find the ASN for IP address a.b.c.d, the DNS query of d.c.b.a.asn.routeviews.org is submitted to DNS system (asn.routeviews.org) and the corresponding ASN and the prefix (100 a.b.c.0/24) can be retrieved in the text field of the DNS answer. As designed, the measurement server is registered in the database beforehand (see Section 2.2.5) with the domain name and IP address mapping pair, as100.dvts.foo.bar and w.x.y.z respectively, so the second mapping can be achieved based on the ordinary DNS lookup.
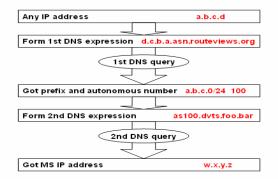

Figure 4 Mapping procedures

2.2.4 Find the measurement servers alone the packet traveling path (auto-discovery)

Finding the measurement servers alone the packet path is most important part for doing the measurement test hop by hop. In SASM, it is assumed that, for identifying the performance segment-by-segment, at least one measurement server should be placed in each autonomous system along the path, as shown in Figure 5.

Since the Internet routing may not be symmetric (the inbound and outbound traffics for a user are not in the same path in opposite direction, as dash line and dotted line for user A shown in Figure 5), measurement tools are designed in such a way that the one way (sending or receiving) and the roundtrip (sending-reflecting-

receiving) tests are both available. Assume that at least one MS has been setup in user's subnet, like MS_A and MS_B in Figure 5. When user A and user B want to setup a DVTS session between them, user A is able to find outgoing path from A to B by taking the ordinary "traceroute" and the incoming path from B to A via the web service issuing remote command "traceroute A" provided by MS_B. The mapping procedures can then be used to find the autonomous system numbers (ASN) and IP addresses of the measurement servers in each AS alone the path, described in Section 2.2.3 (i.e. user A can get the IP addresses of measurement servers MS_4, MS_3, MS_1 and MS_A on its incoming path and measurement servers MS_1, MS_2 MS_4 and MS_B on its outgoing path). Therefore, the network problem can be identified by users and administrators between any two segments along the paths.
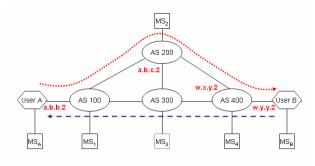

Figure 5 Network topology and measurement servers layout

2.2.5 Measurement server auto-registration

The SASM framework provides a way for newly developed measurement server to do auto-registration in a scalable manner. It includes following steps shown in Figure 6.

(1) Measurement Server Registration

The DNS query method is use to do the registration. The new working MS sends a predefined sequence of DNS query to the registration server (which is also running the DNS service) once an hour. The registration server check the log of the DNS queries, any source IP address matching the predefined sequence is taken as a potential MS. The dvping and additional conformance tests on registration server are then performed periodically with potential measurement server and the source IP address passed the test will be inserted into the database as the measurement server.

(2) Mapping

The mapping procedure described in Section2.2.3 is then used to assign ASN related domain names to the measurement server. If there are more than one servers exiting in the same ASN, the DNS load balancing techniques can be used to retrieve the IP address of the measurement server in a specific autonomous system. Since the DNS has been proven to be a very scalable scheme, it ensures the scalability of this registration process.
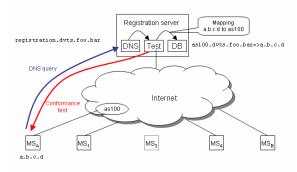

Figure 6 Auto-registration of new measurement server

2.3 SASM Framework

Figure 7 summarizes the general SASM framework. The dvmcast, dvping and DVTS applications described in Section 2.1 working together do the application specific measurement test. The dvbeacon described in Section 2.2.2 ensures the performance inside an autonomous system. The dvtrace described in Section 2.2.4 finds the measurement servers along the packet traveling path and dvreg described in Section 2.2.5 provides a scheme for new measurement server registration. A newly installed measurement server will go through auto-registration process (dvreg) and listed in the database for the corresponding autonomous system. When user A and user B want to run DVTS application, they will do dvtrace to find the measurement servers alone the packet traveling path and enable the measurement test (dvping, etc) with any discovered server at any time to find the network bottlenecks, then solve the problem and have a successful video session.
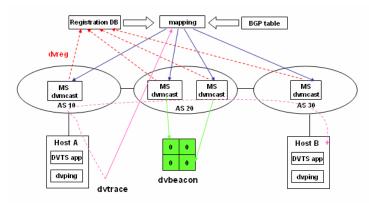

Figure 7 SASM framework

3. CASE STUDY

To demonstrate the effectiveness of SASM framework proposed in this paper, we have developed large-scale distributed measurement servers on national and international academic backbones CERNET2 (AS23910), CERNET (AS4538) and TEIN2 (AS 24489). Several experiments have been done among 25 Pops on CERNET2, among several Pops on CERNET and TEIN2 with DVTS-specific measurement tools installed to support for intra-domain and inter-AS DVTS application collaboration, respectively.

3.1 dvping, dvmcast and dvbeacon

Figure 8 (a) and 8 (b) show the measurement results from DVTS application and the dvping with dvmcast for a pair of nodes under good network condition. Both the end users and network administrators can understand each other concerning the video/audio quality and the network related parameters, such as throughput, the loss and the latency at the same time.

(a)

```
xing@{imni2}[bin]$ dvping 202.112.35.31
ping 202.112.35.31 with dvts data:
send port:8000
recv port:8000
period   :10 s
snd_pkt  rcv_pkt  snd_rate  rcv_rate  loss  rtt
913      913      1Mbps     1Mbps     0%    17.76ms
1824     1824     2Mbps     2Mbps     0%    5.516ms
3646     3646     4Mbps     4Mbps     0%    3.942ms
7291     7291     8Mbps     8Mbps     0%    3.85ms
14578    14578    16Mbps    16Mbps    0%    4.268ms
29153    29153    32Mbps    32Mbps    0%    4.396ms
```

(b)

Figure 8 Result of DVTS-specific network

and application measurement tool

Figure 9 (a) and 9 (b) show the ping testing matrix and dvbeacon for normal ping test (ICMP) and dvping-dvmcast (30Mbps UDP) test from each pair of 25 Pops inside CERNET2 in the debugging time slot, respectively. The horizontal lines are the receiving status while the vertical lines are the sending status of 25 Pops. Though the normal ping test shows the network performance is very good, while the dvping-dvmcast test indicates packet loss in some network links. The comparison of two results proves the necessity of application-specific network tools.
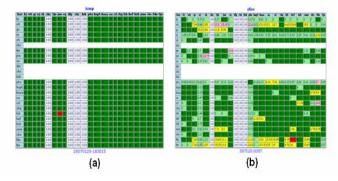

Figure 9 Comparison between ping and dvping

3.2 Measurement Server Auto-discovery

This test is held between a sender in AS24489 and the receiver is in AS 4538. The test topology, the traceroute result with IP address to ASN list mapping and ASN to the measurement server IP address mapping are shown in Figure 10, Figure 11 and Figure 12, respectively. It is clear that at least one measurement server can be obtained using the method proposed in SASM.
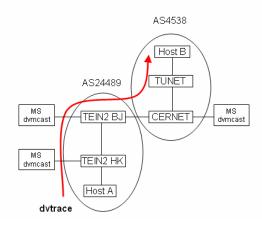

Figure 10. dvtrace test topology

	IP address	1st DNS query expression	ASN
1	202.179.244.1	1.244.179.202.asn.routeviews.org	24489
2	202.179.241.9	9.241.179.202.asn.routeviews.org	24489
3	202.179.241.26	26.241.179.202.asn.routeviews.org	24489
4	202.112.53.17	17.53.112.202.asn.routeviews.org	4538
5	202.112.53.181	181.53.112.202.asn.routeviews.org	4538
6	202.112.38.70	70.38.112.202.asn.routeviews.org	4538
7	59.66.2.26	26.2.66.59.asn.routeviews.org	4538
8	59.66.2.73	73.2.66.59.asn.routeviews.org	4538
9	59.66.2.65	65.2.66.59.asn.routeviews.org	4538
10	59.66.3.105	105.3.66.59.asn.routeviews.org	4538
11	59.66.17.2	2.17.66.59.asn.routeviews.org	4538
12	59.66.17.6	6.17.66.59.asn.routeviews.org	4538

Figure 11 Result of traceroute and the IP to ASN list mapping

# of hops	ASN	2nd DNS query expression	MS IP
3	24489	as24489.test2.foo.bar	202.179.242.102
3	24489	as24489.test2.foo.bar	202.179.244.105
8	4538	as4538.test2.foo.bar	202.112.35.31

Figure 12 Result of ASN list to measurement server list mapping

4. CONCLUSION and FUTURE WORK

In this paper, a new application-specific measurement framework, called SASM, is proposed. Case studies show that key components in SASM are necessary and critical for doing the troubleshooting effectively for high-performance real-time video applications. With the distributed measurement servers running all day-long all over the world, it is convenient not only for end-users to do the test segment-by-segment with or without counterparts at any time, but also for administrators to collect the long term network performance statistics for specific application and

estimate the bandwidth resources availability for certain link. In addition, the scalable measurement server scheme can help the end-users find the test server effectively, especially for inter-autonomous system cases. With SASM, most of the high-performance video application problems could be solved prior to the real multi-sites rehearsal and the heavy negotiation load will be relieved.

The SASM framework can be easily extended for other real-time applications-specific measurement, especially for uncompressed HDTV video (1.2Gbps).

5. REFERENCES

[1] C. Bao, X. Li , Development Trend of Large-scale Distributed Videoconference Environment , *COMPUTER APPLICATION, 2005,Vol24, 853-857*

[2] P.Calyam,W.Mandrawa, M.Shidharan, A.Khan, P.Schopis, H.323 beacon: An H.323 Application Related End-to-End Performance Troubleshooting Tool, *Proceedings of ACM SIGCOMM'04 Workshops, Portland, Oregon,USA, Aug-Sep,2004.*

[3] P.Calyam,W.Mandrawa, M.Shidharan, A.Khan, P.Schopis, Performance measurement and analysis of H.323 Traffic. *PAM Workshop, April, 2004*

[4] Geoff Huston, IPv4-How Long Have we go? *http://www.potaroo.net/ispcol/2003-08/ale.html*

[5] K.kobayashi, A Ogawa, S.Casner, C.Bormann, RTP Payload Format for DV (IEC6183) Video. *[RFC 3189]*

[6] K.kobayashi, A Ogawa, S.Casner, C.Bormann, RTP Payload Format for 12-bitDAT Audio and 24-bit Linear Sampled Audio. *[RFC 3190]*

[7] M. Kutzko, T. Rimor\vsky, J. Ectabrook, J. Dugan, Multicast Beacon. *http://dast.nlanr.net/projects/Beacon/*

[8] M.Luckie, J.Mcgregor, H.Braun, Towards Improving Packet Probing Techniques, *Proceedings of ACM SIGCOMM'01 , Workshops, 2001*

[9] M.Luckie, J.Mcgregor , IPMP: IP Measurement Protocol. *http://moat.nlanr.net/Papers/PAM02-IPMP.pdf*

[10] N.Nakashima, K. Okamura, Telemedicine with Digital Video Transport System in Asia-Pacific area. *Proceedings of the 19th International conference on Advanced Information Networking and Application, Vol.(253-257),2005*

[11] A Tirumala, F. Qin, J.Dugan, etc. Iperf , *http://dast.nlanr.net/projects/Iperf/#whatis*

[12] P.Yalagandula, M. Dahlin, A Scalable Distributed Information management System, *Proceedings of ACM SIGCOMM'04 Portland, Oregon,USA, Aug-Sep,2004.*

[13] P.Yalagandula, P.Sharma, S.Banerjee, S.Basu, J.Lee, S3: A Scalable Sensing Service for Monitoring Large Networked System, *Proceedings of SIGCOMM'06, Workshops, Pisa,Italy, Sep,2006*

[14] China Education and Research Network (CERNET) *http://www.edu.cn/english_1369/index.shtml*

[15] Next Generation of China Education and Research Network (CERNET2) *http://www.cernet2.edu.cn*

[16] DV Stream on IEEE1394 Encapsulated into IP. *http://www.sfc.wide.ad.jp/DVTS/*

[17] NLANR: Active Measurment Project, http://amp.nlanr.net/

[18] Trans-Eurasia Information Network (TEIN2) *http://www.TEIN2.net*

[19] University of Oregon Route Views Project, *http://www.routeviews.org/*

[20] Windows XP Professional Product Documentation *http://www.microsoft.com/resources/documentation/windows/xp/*