Measurement and Management of Internet Services

C. Darst, S. Ramanathan
Hewlett Packard
3404 E. Harmony Road
Fort Collins, CO
USA
{chuck_darst, srinivas_ramanathan} @hp.com

Abstract

It is increasingly important for Internet Service Providers (ISPs) to monitor their systems to detect problems in advance of customer complaints. Accurate, yet rapid, diagnosis of detected problems is essential for minimizing customer-perceived service downtimes. Such requirements warrant a new generation of technologies for service quality measurement and problem diagnosis.

End-to-end performance management requires analysis of the ISP network, servers, applications, and infrastructure components. This paper proposes a methodology and measurement instrumentation for managing ISP service quality. This paper presents case studies that demonstrate the use of performance measurements for problem detection and isolation. The results, which highlight the practical utility of the measurements, have been obtained from real-world measurements in commercial ISP environments and the general Internet.

Keywords

QOS, Internet Services, Performance Monitoring, Measurement, Service Level Management, Firehunter

1. Introduction

Quality of Service (QOS) is a broadly discussed topic within the Internet community. Typically, QOS is discussed in relation to network traffic flows and associated prioritization schemes. At the same time, significant attention is being paid to differentiated and tiered Internet provided services such as web hosting and electronic commerce, and associated service-level management. One more emerging topic revolves around performance monitoring and management. This paper spans these areas analyzing QOS with respect to the availability and performance of Internet services.

Service oriented QOS can be a key differentiator for Internet Service Providers (ISPs). In a PC Week study (Wetsel 1997)**,** 96 percent of respondents listed service availability as their main consideration in choosing ISPs. Furthermore, 93 percent of

respondents perceived service performance as their second key expectation from ISPs. Measuring and monitoring of provided services is also a fundamental building block for commercial offering of tiered services.

To meet the expectations of customers and to help attract new ones, ISPs need to measure and manage their QOS. This requires a shift in the management paradigm for many ISPs. From monitoring IP port availability and managing network links in terms of metrics such as packet loss and delay, ISP operations personnel must look at a new set of metrics that represent QOS from a user perspective. Examples include the accessibility of a specific Usenet newsgroup, the delay between transmission and reception of an email message, and web page downloading time.

More than ever before, it is becoming increasingly important for ISPs to monitor their systems to detect problems in advance of customer complaints. Furthermore, accurate yet rapid diagnosis of detected problems is essential to minimize customer-perceived service downtimes. All of these requirements warrant a new generation of technologies for service quality measurement and problem diagnosis.

End-to-end performance management requires analysis of the ISP network, servers, applications, and infrastructure components. In this paper, HP Firehunter is used to illustrate concepts associated with measuring the performance of Internet services such as email, news, web, and network access. ISPs use HP Firehunter to measure, manage, and report the quality of their services.

This paper describes proposes a methodology and integrated instrumentation for managing quality of service for ISPs. Section 2 describes measurement technologies. Section 3 discusses details of the measurements employed. Section 4 presents case studies that demonstrate using these measurements for problem detection and isolation. Section 5 presents related work. Our conclusions are presented in Section 6. The results, which highlight the practical utility of all of the measurements, have been obtained from a real-world deployment of the measurements in commercial ISP environments.

2. Measurement Technologies

Two main classes of measurements, active and passive, are discussed in this section.

2.1 Active Measurements

Active measurements simulate actual user transactions such as sending an email message. They explicitly stimulate traffic to networks, servers, and service applications to assess the elements. For example, to measure the availability and performance of a web server, a test actively measures an HTTP GET request for a web page and observes the status code in the response from the server as well as the total response time for retrieving the web page.

To truly reflect customer-perceived problems, the active measurements should use the same network path and the same set of infrastructure services (DNS, Network File Service (NFS)) that customers use. If this is not the case, additional measurements may be needed and correlated to assess the state of any elements of interest. For instance, if the performance of a DNS server is assessed using a different

path than the one customers use, then the DNS response time should be correlated with a customer-oriented network path response time. In this case, the DNS server's performance is poor only if the DNS response time is abnormally high and the network response time is not.

Active measurements do not rely on any monitoring capabilities built into the networks, servers, and service applications. The measurements can be initiated from multiple locations providing closer approximation of customer's perceived performance. If they are not carefully engineered, however, these measurements can end up impacting the QOS they are intended to measure since they introduce additional traffic!

Active measurement can be taken (externally) across ISPs without requiring access to an ISP's systems. Such performance information typically does not include associated system or networking data limiting detailed problem diagnosis and capacity planning. Consequently, the utility of external measurements might be limited to service quality assessment and preliminary problem diagnosis.

2.2 Passive Measurements

Passive measurements collect local system information. Generally, measurements in this class utilize instrumentation built into network, server, and service application components. They do not require generation of traffic (beyond information collection), and have minimal impact on ISP systems. Besides being useful for service quality assessment, passive measurements can also provide information about resource and service usage, which is critical for detailed problem diagnosis and capacity planning.

Resource and service usage information can be obtained from the servers. This includes data such as memory and CPU utilization. Networking and interface port statistics provides additional information related to system loads. This class of information is useful in diagnosing causes of overall service quality problems. Additionally, trend information can be analyzed over time to predict potential system bottlenecks.

Log files and management information bases (MIBs) provide additional information in monitoring service-oriented QOS. TCP-based Web, email, and news services involve explicit communication of acknowledgements between customer client applications (for example, web browsers, and email clients) servers. Based on acknowledgement feedback, service applications can detect both service transaction start and end. As part of normal operation, the service applications can measure the service performance in terms of response times for different service transactions. This information can be stored in log files or MIBs for further analysis by management applications. External network probes can be used to snoop network transmissions and analyze the captured packets to yield measurements of services and networks.

2.3 An Architecture for the Use Of Active And Passive Measurements

Figure 1 illustrates the components of a system such as Firehunter. A diagnostic measurement server (DMS) collects and processes data collected by agents that make

active measurements of service quality. Software agents operating on the ISP servers track resource and service usage passively. The combination of active and passive measurements provides information that is critical for effective operational monitoring and capacity planning for ISPs. To maximize the effectiveness of the measurements and to minimize their impact on the ISP system, a flexible service model allows the user to easily configure where the agents are placed and how often the measurements are executed.

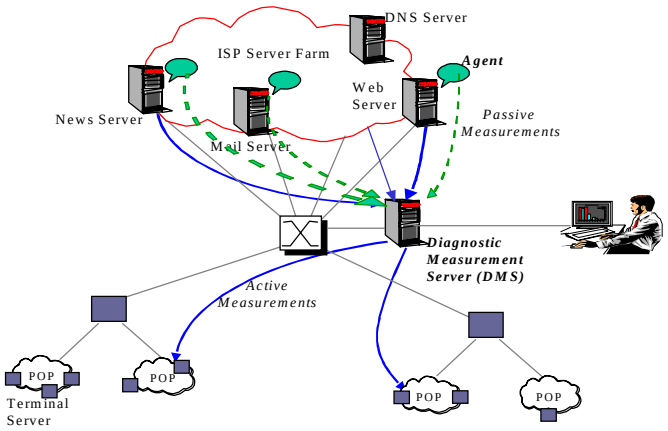


Figure 1. Components of Firehunter: the deployment of a DMS and measurement agents in an ISP server farm

Active measurements can also be taken remotely - from the DMS. For example, a remote agent could be installed in one or more POPs to monitor web server (http) performance from the remote locations. The measurement data is then uploaded to the central DMS for aggregation into the complete service model.

3. Measurement Details

Systems such as Firehunter provide service quality, network, infrastructure service, and server measurements. This section describes these measurements in detail.

3.1 Service Quality Measurements

To measure the QOS of web, email, and news services, Firehunter includes tests:

Web service measurements: The monitoring test for web services uses active measurements to assess the web services' availability and performance. Like a typical web client retrieving a web page, this web service test resolves the IP address of a target web server, establishes a TCP connection with the web server, and then issues a GET request to a specified web page). By interpreting the HTTP response header

returned by the server, the web service test determines the availability of the web service. The overall response time to retrieve the web page is determined. The test derives a breakdown of the overall response time into the following categories: DNS resolution time, TCP connection establishment time, server response time (the time between when the GET request is issued to the time when the HTTP response header is received), and Data transfer time (the time between when the HTTP response header is received and the web page retrieval is completed).
Individual response time components can reveal different potential bottlenecks. An increase in total delay or in the DNS response time is a typical indicator of DNS-related problems. A large TCP connection establishment time is often indicative of bottlenecks in the server host (note that although the TCP connection is established to the port the web server application is listening to, TCP connection establishment is entirely handled in the kernel, rather than in the web server application). The server response time is indicative of the web server application processing delays.

Email service measurements: The ability of a customer to send mail to other users (whether they are customers of the same ISP or of a different ISP) is distinct from the ability of the customer to receive mail from the email servers. To assess the availability and performance of the distinct email operations, the following active measurement tests are used:
Mail delivery test: assesses the capabilities of an ISP's email system to accept mail messages for delivery to one or more target destinations. ISPs may use different mail servers to handle messages destined for local customers and those destined for external locations on the Internet; consequently, different local and remote destinations can be specified as targets to the mail delivery test. This test uses the Simple Mail Transfer Protocol (SMTP) to communicate with one or more ISP mail servers, transmits the specified mail messages, and reports the availability and performance of the mail servers.
Mail retrieval test: emulates a customer accessing a mailbox from the ISP's mail server. While doing so, this test assesses the availability and performance of the mail retrieval components of the email system independent of the mail delivery components. In addition, this test can also measure the delay between the transmission of a mail message and its reception at the intended destination. Figure 2 depicts the operation of this test, which uses the Post Office Protocol Version 3 (POP3) for communication with an ISP's POP3-based mail server.
To obtain consistent measures over time, a constant sized mailbox should be used for retrieval. Like the web service test, the mail retrieval test tracks the overall response time for accessing the mailbox as well as the individual components of response time: TCP connection establishment time, time to authenticate the customer, and time to retrieve the customer's mailbox. These response time components can be used to identify potential bottlenecks that may exist during retrieval of mail messages in the ISP system.

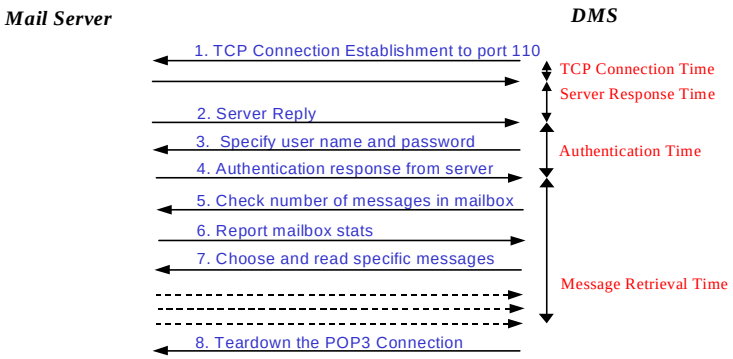


Figure 2. Operation of the mail retrieval measurement test

While retrieving messages from a mailbox, the mail retrieval test can be instructed to perform additional actions for these messages. For instance, the test can be instructed to estimate the delay between the transmission and reception of the message by comparing the origination time of the mail message (included in the mail header) and the retrieval time of the message.

News service measurements: Availability and performance of the news service is measured using a news service test. The operation of this test is very similar to that of the other active measurement tests described earlier.

3.2 Network Measurements

As customers typically use different network paths to access web, email, and news services from the server farm, the QOS observed by customers is dependent on the POP sites they use to connect to the ISP. To obtain an end-to-end perspective of service quality, the service quality measurements described above must be complemented with measurements of the network's effects on service quality.

ISPs have traditionally used tools such as *ping* and *traceroute* to detect and diagnose network problems. Application servers, routers, and terminal servers in the different POPs can be targets for the network connectivity and delay measurements.

While *ping* and *traceroute* give an indication of changes in network characteristics, these tools are not sufficient to quantify the impact of these changes on the availability and performance of customer-visible services. For services that use TCP for reliable communication, a useful network performance metric is *throughput*, defined as the rate of reliable transmission of packets between a source and a destination. The throughput achievable between any source-destination pair is a function of several factors such as the socket buffer size in use, characteristics of the source and destination's TCP implementation, processing capabilities of the source and destination, bursts of packet loss (if any), round-trip packet transmission delays, and so forth. The complexity of this relationship makes it almost impossible to estimate the throughput achievable based on packet loss and delay measurements available from *ping*.

Many public domain tools such as *throughput TCP (ttcp)* and *netperf* (Jones 1996) have been widely used for measuring throughput. A key drawback of these tools is the need for custom software applications to be executed at the source and destination, respectively, in order to enable the measurement. To enable throughput measurements without requiring special-purpose instrumentation at the targets, Firehunter's throughput monitoring test builds on the concept of the *Traceroute Reno (Treno)* tool that has been developed as part of the Internet Provider Performance Metrics (IPPM) working group of the Internet Engineering Task Force (Mathis and Mahdavi 1996). *Treno* emulates a TCP-based data transfer using User Datagram Protocol (UDP) packets to transmit data to a target. In its basic form, *Treno* was intended to permit customers to compare the performance offered by different network providers, and for network providers to monitor the performance of their networks.

Firehunter's throughput monitoring test modifies the concepts of by restricting the amount of data transferred during each measurement to match typical data transfer sizes that customers use when retrieving web, email, and news content., Firehunter's throughput monitoring test ensures that its measurements reflect customer perceptions of network throughput. Since it does not rely on custom software in the targets for throughput measurements, Firehunter's throughput monitoring test can measure throughput to any IP-capable device, such as application servers, routers, terminal servers in the POP sites, or even customer PCs and terminals. Firehunter's throughput monitoring and packet loss tests have been shown to produce consistent (within 97%) results compared to other similar tools that require both source and destination measuring components.

An approximation of end-to-end service quality can be obtained by combining the web, email, and news service quality measurements with throughput measurements directed to customer PCs and terminals. In the event that a reduction in network throughput is detected, *ping*, *traceroute*, and other tools can be used to further isolate the root cause of the problem.

3.3 Infrastructure Service Measurements

Although they are not directly visible to customers, infrastructure services such as DNS, which handles translation of host names to IP addresses, and the NFS, which enables remote data access, play a crucial role in the operation of Internet services. Any deterioration in availability and performance of these infrastructure services usually manifests as customer-visible problems with one or more of the Internet services.

DNS measurements: Truly accurate measurements that assess DNS performance as perceived by customers can only be obtained using passive measurements that track address resolutions requested by customer applications. Such measurements can be made either at the DNS server (requiring modifications to the server application), at DNS clients (for example, some web proxy servers log DNS response time as one of the components of the overall response time for customer requests), or using non-intrusive external probes. In the absence of such passive monitoring capabilities in most ISP systems, Firehunter uses active measurements to emulate typical DNS requests and assess DNS availability and performance.

Since the ability of a DNS server to service requests from its cache is independent of its ability to resolve mappings not in the cache, the DNS test includes separate measures of DNS cache hits and misses for a server. By issuing a non-recursive query to a DNS server, the test forces the server to respond based on its local state alone, thereby emulating a DNS cache hit at the server. By requesting address resolution to a well-known Internet site and by observing the status returned by the DNS server in its response, the DNS test measures the availability of the DNS server. To measure the performance of DNS cache misses, the DNS test issues a request for address resolution for a randomly chosen host. Comparison of DNS response times with the network response times provides an indication of whether the network or the DNS server is a performance bottleneck.

NFS measurements[1]: One of the common ways of scaling an ISP system is by deploying back-end content servers that store web, email, and news content, and permit front-end servers (FESs) to access the content via NFS. Passive measurements that track NFS statistics maintained by the server and client machines yield various statistics about the NFS sub-system. On the server end, the number of NFS calls handled by the server over time can be an indicator of potential overload conditions. On the client end, the rate of NFS calls issued by the client, the percentage of duplicate responses received from the server(s), the percentage of time-outs observed by the client, and the percentage of retransmissions can all be tracked over time to detect NFS anomalies.

Comparing the percentage of retransmissions with the duplicate responses provides insight into potential problems. If few duplicate responses have been received, then the interconnecting network could be causing problems. If the percentage of duplicate responses is almost equal in magnitude to the percentage of retransmissions, this implicates a slow-down at the server end as the cause of the problem (for example, because there are too few NFS server (*nfsd)* processes executing on the server). A look at the client statistics that indicate the number of NFS requests waiting for service can indicate if there is a bottleneck at the client end. When there are too few NFS client (*biod)* processes executing on the client to handle remote NFS accesses, the count of waiting requests grows.

3.4 Server Measurements

Servers can yield a wealth of information that is critically important for isolating problems. Firehunter gathers measurements for the CPU utilization of the server, free memory availability, virtual memory page scanning rate, and packet transmission and reception rate through each of the interfaces of the server. For servers that support TCP-based services (web, email, news, and so forth), the rate of connections to and from the server and the number of connections currently established on the server are useful measures of the server's workload. A breakdown of connections based on the specific ports to which they relate is also available, thereby enabling workload to be determined for each service (SMTP, POP3, HTTP, and so forth), even in cases when

[1] NFS measurements are not currently available in Firehunter. Future releases are expected to include NFS measurement capabilities.

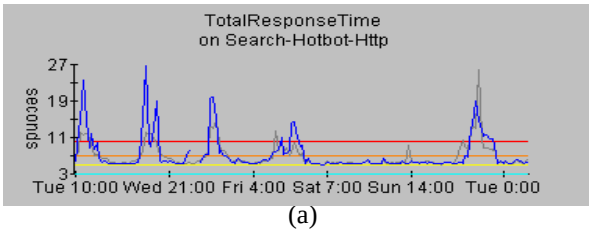
the same physical host supports multiple services. By monitoring the different states of the TCP connections to a server, it is possible to detect abnormal conditions such as the malfunctioning of specific servers or the occurrence of SYN attacks.

4. Demonstration of Performance-based Monitoring

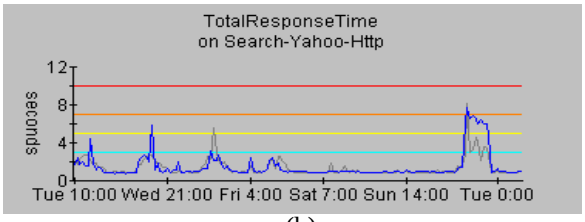
This section presents examples of how service level, performance-based measurement and monitoring can quickly detect and isolate problems. The examples are based on using Firehunter measurements in real-world ISP environments.

4.1 DNS Problem

Since DNS is a common infrastructure service that impacts all the ISP services, such as web, email, and news, failure of DNS is often widely noticed. While hard failures caused by DNS availability problems are easily spotted, soft failures caused by DNS performance problems often take a long time to resolve. The DNS measurements can be deployed for spotting both cases.



(a)



(b)

Figure 3. Internet Search Engine Response Times

The two Firehunter graphs (Figure 3) show the performance monitoring of two Internet search engine web sites over a week time period. The test being monitored includes requesting the search engines to complete a query on "Firehunter". Clearly visible towards the end of the second graph (from yahoo) is a five-second delay that occurred on the Monday. This type of performance problem could easy be misdiagnosed as a system or network problem.

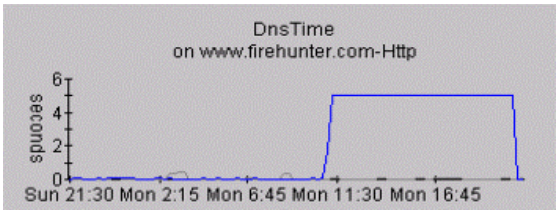


Figure 4. DNS Response Times

As Firehunter also tracks the individual components of web page retrievals, more detailed performance graphs (Figure 4) quickly showed that it was taking five seconds to resolve DNS lookups. The five-second DNS delay was observed for all sites tracked during the time of the test. In Figures 3 and 4, the five-second delay is seen impacting Hotbot, Yahoo, and the Firehunter web sites. This was tracked back to a failure of the primary DNS server. Five seconds was the timeout before the secondary DNS server then handled the request.

Performance monitoring of DNS in addition to the higher level services illustrates the impact of supporting services on total response time of customer facing services. Measuring the component times of service transactions also allows for faster fault isolation and correction.

4.2 News Problem Diagnosis

Based on ongoing measurements of an ISP's news service, we recorded typical response times for retrieving a set of news articles (10 Kb of data) in the range of 1 to 6 seconds. In the scenario under consideration, it was observed that the response times for news access dramatically increased to over 100 seconds at certain times of the day. Further analysis of the measurement results (see Figure 5) revealed that one of three FESs, news-FES3, was performing significantly worse than the others.

The poor performance continued for several days, as the ISP's operations personnel failed to detect that the problem existed, let alone fix the problem. A week later, the performance measured for all three FESs had degraded significantly. While news-FES3 performed poorly in the period 8/19-8/21, all three FESs experienced problems on 8/22.

Analysis of the breakdown of the FESs response time indicated that on 8/22, while the set-up time dominated during the period from 7 AM to 10 AM, the server response time was the dominant component during the period from 10 AM onwards. The changes in the dominant response time components suggest that there were two different problems (with potentially different sources) affecting news performance. A similar behavior was observed for the other FESs as well, indicating that the same problems were affecting all the FESs.

Passive measurements available from agents operating on one of the FESs, news-FES1, provided additional information essential for diagnosis. Comparison of the rates of TCP connections handled by news-FES1 on 8/22 with baselines computed from historical data indicated that there had not been any increase in the number of connections to and from the news-FES. Analysis of the rates of packets to and from news-FES1's external interface also failed to indicate any change in customer-

generated news workload. However, analysis of the rates of packets to and from news-FES1's internal interface revealed the source of the problem.

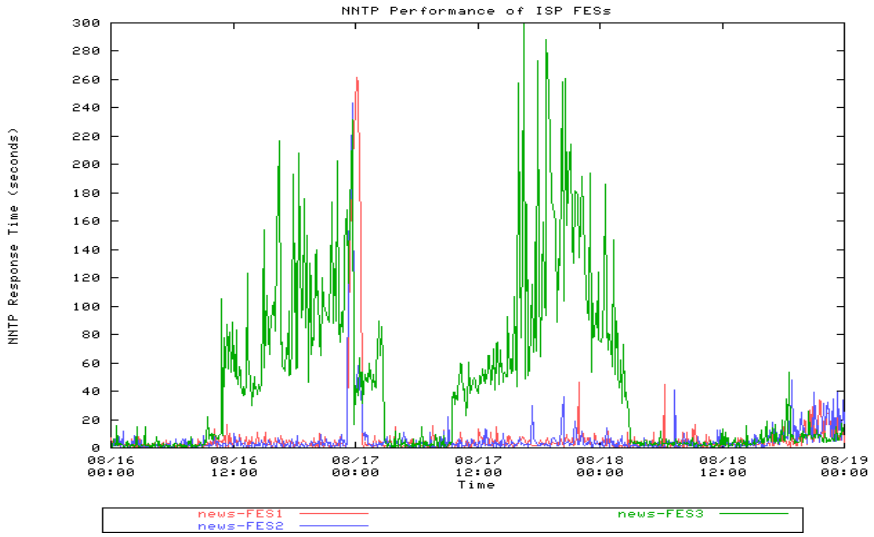


Figure 5. News FES Performance

Figure 6 depicts the packet traffic handled by news-FES1's internal interface over a week. There were two traffic peaks—one on 8/17 and the other on 8/22—corresponding to an incoming traffic rate of over 45,000 packet/minute, which corresponded to times when news performance had degraded. The second peak coincides with the 7 AM to 10 AM time period on 8/22. The above analysis found that the source of the first problem is congestion on news-FES1's network interface. Rerouting of traffic from the 10Mbps interface to a 100Mbps interface around this time removed the performance bottleneck.

The second problem observed on 8/22 from 10 AM onwards was attributable to NFS problems. While there had not been a detectable increase in NFS calls per minute issued by news-FES1 to the NFS server, the percentage of NFS time-outs had increased dramatically from close to 0 to almost 10 percent. Comparison of network response times between news-FES1 and the NFS server did not indicate any change from normal. This leads to the conclusion that a slowdown of the NFS server was the cause of the problem. Since the NFS server is accessed by all the news FESs, performance problems were observed when accessing all of the news FESs.

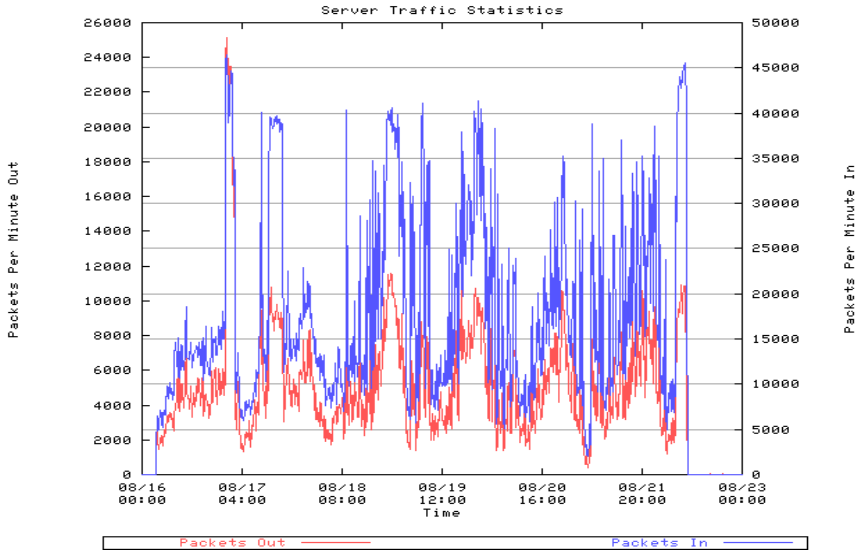


Figure 6. Packet traffic handled by the internal interface of news-FES

4.3 Email Performance Tuning

Another example dealt with the performance of an ISP's mail delivery system. In this case, we tracked the delay between the transmission of a message and its receipt at a local customer mail account. Starting with a mail delivery delay of about 15 minutes on 8/27, mail delivery performance of the ISP's mail system continued to degrade. By 9/10, mail delivery delay had reached close to 40 minutes at peak times. This rapid increase in mail delivery delays was attributable to a failure of the mail queue processing algorithms implemented by the mail server application. As is noticeable from the figure, tuning of the mail server application, together with the introduction of new mail handling policies by the ISP (for example, to queue mail messages for no more than 3 days), resulted in a significant drop in the mail delivery time. By 9/22 mail delivery delays were well under 5 minutes. Besides providing measurements that can enable ISPs to assess the quality of services they offer, Firehunter provided a means by which the ISP operation staff could directly observe the impact of the performance tuning and system reconfigurations they perform.

4.4 Network Monitoring

The Firehunter web site (www.firehunter.com) monitors a number of different well known web sites and services including www.whitehouse.gov for the United States White House. Traffic, availability, and response times of this site have historically been highly sensitive to breaking news stories that impact the White House. Figure 7 is a network percentage packet loss performance graph for the week of September 10, 1998 through September 17, 1998. Friday, September 11, was the date the Starr report was issued. While the White House did not post a copy of the report, it did

post rebuttal information and experienced significant availability and performance problems during the time in question.

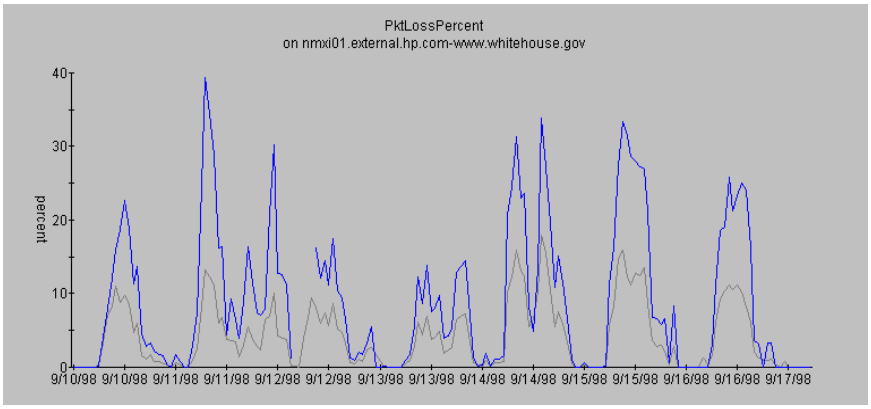


Figure 7. www.whitehouse.gov Network Packet Loss

Items of interest related to Figure 7 include the very high percentage of network packet loss compared to typical days for this site and other web sites normally less than 10%. The heavy network load was also reflected in the associated network throughput statistics tracked for this site (not shown) and overall availability and response times.

5. Relation to Existing Products and Tools

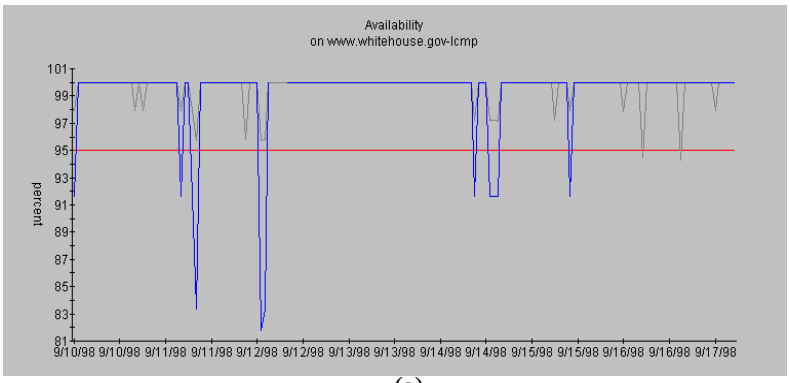
Conventional network and system management systems provide limited capabilities to assist an ISP in detection and diagnosis of problems relating to Internet services. Over the years, several measurement tools that have been developed and made available in the public domain for monitoring IP networks and servers. Tools such as *ping* and *traceroute* have been used for network monitoring and diagnosis, and those such as *top*, *vmstat*, and *iostat* have been used for monitoring servers. The lack of an integrated, easy-to-use monitoring system has meant that ISP operations personnel spend a significant amount of time manually troubleshooting problems. The time taken for problem diagnosis has a direct impact on customer-perceived QOS. As the complexity of ISP systems grows to handle scale and to deliver newer, more sophisticated services, problem detection and diagnosis have tended to take longer and longer. This has resulted in growing numbers of unsatisfied customers, contributing to an industry-wide churn (Wetzel 1997).

Some system-oriented toolkits have attempted to simplify the task of problem diagnosis for ISPs by providing a variety of statistics relating to the performance of their servers. However, such toolkits do not provide any capability for modeling and measuring Internet services. The focus on management of services, rather than management of just servers, is a key distinguishing characteristics of a service-level, performance-monitoring system.

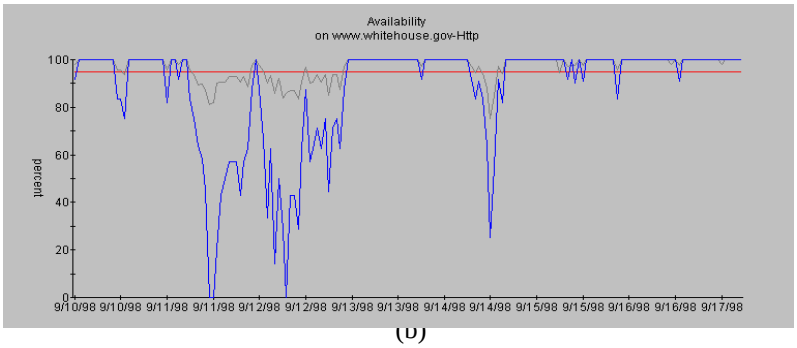
Pinging tools are also frequently used for checking server availability. Port-checking methods attempt to establish connections to respective TCP ports and issue a single

service-specific command - not truly representing customer accesses to the ISP services. Consequently, these measurements are not sufficient to assess the service quality of ISP services. For example, these measurements do not indicate how long it will take a customer to download a 10 Kb set of news articles from a news server. Moreover, since it predominantly uses active measurements, these tools do not provide sufficient information for problem diagnosis.

Figure 8 illustrates importance differences when checking the availability of a web site between ping and the actual downloading of a web page. The graphs in are from the White House web site for the week after the release of the Starr report.



(a)



(b)

Figure 8. www.whitehouse.gov Availability Comparison

The heavy network and system load during this time frame caused severe problems in availability and response time performance. Ping is typically handled at the system kernel level, and does show some availability problems for the White House web site in graph (a). However, this only has minimal correlation with a user's ability to actually download web pages from a web site as shown in the second graph (b) in Figure 8.

Other items of note related to the graphs in Figure 8: gaps in the measurement data recorded correspond to periods of (www.whitehouse.com) unavailability. A purpose of the above comparison is to show that ping and IP-port availability checking tools and methods, by themselves, are insufficient in detection and isolation of many

Internet service problems. Ping is a simple and fast test often providing the first indication of a service problem; consequently, a ping test is included in Firehunter.
One early system that attempted to provide an integrated solution for ISPs is the network operations console, NOCOL (Enger and Reynolds 1993). Firehunter uses a manager-agent architecture similar to NOCOL's. In another approach, Inverse Networks devised a measurement suite for relative comparison of the QOS delivered for dial-up access services offered by different ISPs (Poole and Doan 1997). These measurements, are designed to emulate typical customer accesses, and are executed from one or more locations. By dialing in to each of the ISP's POP sites periodically, the measurements provide an estimate of the percentage of times customers are likely to experience busy signals while connecting to the ISP.
The IPPM working group of the Internet Engineering Task Force has been exploring tools for assessing network quality (Paxson et al. 1998). Initial draft metrics have been proposed, and early prototypes of tools are being evaluated. As the metrics become standard, future releases of Firehunter will include appropriate IPPM measurements in their measurement suites.

7. Conclusions

This document has described performance-based measurements of a service-oriented solution for the measurement and management of ISP services. Using a combination of active and passive measurements of ISP services, infrastructure, networks, and servers, performance-based monitoring gathers information critical for operational monitoring, capacity planning, and customer support.

References

1. Brian Atkins, HP OpenView Firehunter Service Models and Baselining, Hewlett-Packard Company, March 1998.

2. Mike Avery, WhatsUp Gold enhances network supervision tasks, *InfoWorld*, Vol. 20, Issue4, January 26, 1998.

3. Adrian Cockcroft, Advanced monitoring and tuning, *Sun World On-Line*, October 1995, http://www.sun.com/951001/columns/adrian/column2.html

4. R. Enger and J. Reynolds, RFC 1470: FYI on a network management tool catalog: Tools for monitoring and debugging TCP/IP Internets and interconnected devices, June 1993.

5. Rick Jones, Netperf: A network performance benchmark: Revision 2.1, Hewlett-Packard Company, February 1996.

6. Matthew Mathis and Jamshid Mahdavi, Diagnosing Internet congestion with a transport layer performance tool, *Proceedings of INET'96*, June 1996, http://www.psc.edu/~mathis/htmlpapers/inet96.treno.html

7. Vern Paxson, Guy Almes, Jamshid Mahdavi, and Matthew Mathis, Framework for IP Performance Metrics, Internet Engineering Task Force (IETF) Network working group Internet Draft, February 1998.

8. Jackie Poole and Amy Doan, Inverse profile checks up on ISP performance, InfoWorld, April 28, 1997: 68.

9. Rebecca Wetzel, Customers rate ISP services, *PC Week*, November 1997, http://www.zdnet.com/pcweek/sr/1110/10isp.html

Biography

Chuck Darst is the HP Firehunter Product Manager in Fort Collins, Colorado. His career has been oriented around the monitoring and managing of communication networks and associated services - spanning local area networks to the public switched network. He received a Bachelor of Science degree in Electrical Engineering from Colorado State University, Fort Collins, and a Master of Science degree in Computer Science from the University of Illinois, Champaign-Urbana.

Srinivas Ramanathan is a research scientist at Hewlett-Packard Laboratories, Palo Alto, where his research focuses on measurement and management for emerging broadband networks, Internet technologies, and services. Srinivas received the B.Tech degree in Chemical Engineering from Anna University, Madras, India in 1988, the M.Tech degree in Computer Science and Engineering from the Indian Institute of Technology, Madras, India in 1990, and the Ph.D. degree in Computer Science and Engineering from the University of California, San Diego, in 1994.