Standardized Active Measurements on a Tier 1 IP Backbone

Leonard Ciavattone, Alfred Morton, and Gomathi Ramachandran, AT&T Laboratories

ABSTRACT

Synthetic or active measurements are often used to characterize IP performance; however, it is rare to find them used to resolve problems in an operational setting. In this article we show that the active monitoring system in the AT&T IP backbone provides a comprehensive view of network performance that is complementary to traditional element level monitoring, making it an integral part of network management. This paper discusses the design and implementation of these active measurements in the network. We continuously monitor "path-level" performance metrics such as round-trip delay, loss, jitter, and reordering events to proactively detect impairments. Our system relies on the promotion of key metrics to the operational displays, while maintaining a rich set of statistics for analyzing rare and unforeseen events. This timely information enables us to react quickly to performance degradation, avoiding any sustained effect on customer applications. The results also help us understand the network's ability to support time-sensitive application performance. Selected "interesting" events observed are presented, including detection of degradation caused by low-level bit errors on a physical link, detection of route changes on the network and their impact on real-time applications, and finally detection of reordering caused by forwarding loops.

ACTIVE MEASUREMENTS AND NETWORK OPERATIONS

Numerous studies have used active measurements on IP networks to characterize performance and isolate interesting events [1, 2], and to understand the statistical properties of Internet packet transport [3]. Research applications of network monitoring may be tailored to extract specific properties of interest, and may need several offline processing steps to get the desired results. In contrast, operational measurements have different goals. They must:

- Capture all customer-affecting events (especially those with well-known characteristics)
- Be easily understood by network technicians
- Have alerts or alarms that are unambiguous and lead to corrective actions
- Minimize false positives
- Provide near real-time status and notification
- Complement a traditional fault/passive management system

Thus, operational active measurements benefit from the research work, but have their own challenges and limitations.

In this article we will describe a network-wide measurement system that is currently in operation on a tier 1 Internet service provider (ISP) backbone. This measurement system continuously monitors the path-level performance of the backbone, and is designed to produce operational level alerts and data, providing the basis for proactive issue resolution.

BACKGROUND: TRADITIONAL NETWORK OPERATIONS MEASUREMENTS

Traditional network management measurements have relied heavily on element-level statistics and alerts. These measurements include fault alarms (traps) or counter-based element measurements (e.g., inbound octets on an interface).

Traditional fault alarms are often triggered by equipment failures. In the traditional IP environment, network layer recovery was thought to be sufficient (along with redundant network design). However, real-time applications can be severely affected by outages of 2–10 s, while network layer recovery often takes 10–15 s (see "Results"). Thus, there is a need for proactive notification of degradation in addition to element failures.

Passive measurements monitor the performance seen at a single network element (interface, router, link). They collect link-utilization, router load, errors, queue drops, and so on. These measurements are essential for network management and cannot be replaced by active measurements. Moreover, monitoring link utilization and trending the rate of its increase is an effective metric for managing congestion by link. On the other hand, the end-to-end path performance of a packet (delay, delay variation, or

0163-6804/03/$17.00 © 2003 IEEE

loss) cannot be measured passively at an arbitrary single point. Using passive measurements to get an estimate of end-to-end performance would require:

- Prior knowledge of the path
- Synchronized collection
- Knowledge of the measurement's relationship to the end-to-end path (e.g., delay may be somewhat additive over segments, loss will be a conditional probability)

As these three requirements are difficult to meet and may be beyond the reach of current technology, active measurements are best suited to estimate delay, loss, and delay variation over path segments.

In this article we describe our active measurement system and its use in detail. In the following section we discuss the needs that motivate these new measurements. We outline the principles of our measurement design, including objectives, test design, and deployment. We describe the metrics we collect, along with useful ways to summarize the results. We list some of the issues with our measurement system. We present detailed results, including observations of low-level loss, excessive delay variation, route changes, and route loops. A later section contains a short summary of other types of active measurement systems (and their issues). Finally, we summarize our conclusions from this work.

Measurement Motivation: Fast Detection of Application Performance Events

The ability to correctly estimate packet transfer performance during congestion/recovery events is especially valuable for assessing the performance of real-time applications. During a voice over IP (VoIP) phone call, users experience the network's packet transfer performance continuously for minutes at a time, and even short intervals of poor performance lead to dissatisfaction. In contrast, Web users seldom sample the performance for more than tens of seconds while requesting HTML pages, and file transfer users usually turn their attention to other activities while a multiminute transfer is in progress. Both Web and ftp users can easily accommodate short periods of congestion/poor performance. Thus, in the subsections below we summarize the key metrics for a variety of applications. We note that International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Recommendation Y.1541 provides a small set of classes with objectives that satisfy the needs of these applications. The need to meet such objectives motivates much of our need for active measurement monitoring.

Measurement Design

An active measurement system has been operating on the AT&T IP backbone since 1998. The original system was designed to:

- Estimate network performance, primarily in terms of delay between networking centers
- Publish timely performance information for the benefit of current and potential customers

In practice, the tool was also used for some troubleshooting, but limitations in packet rates made it difficult to identify and characterize root causes. Furthermore, the set of customer applications (e.g., ftp, SMTP, telnet) commonly used at the time was forgiving and well adapted to the challenges of IP networking. We recognized the value of active measurements in diagnosing network problems. Thus, the current measurement design includes both trouble cause analysis and the needs of a more demanding set of applications (e.g., VoIP, HTTP/Web, video).

Objectives

The design of active measurements responds to these main considerations:

- Practical limitations of the measurement architecture (desire to use a standard/unmodified UNIX kernel, expense of servers and server deployment in networking centers, difficulty in acquiring a GPS feed, amount of data generated).
- The added demands of real-time applications on network performance. In addition to packet loss and packet transfer delay limits, real-time traffic requires consistent packet transfer delay. There is also a need for a probe to detect network delay variation, during both exception congestion events and the network's normal operation.
- Proactive detection of intervals with performance below desired levels in time to take corrective action.

Addressing these considerations in a measurement system will reveal both the steady performance of the network and variability of performance. There is no intent here to replace passive measurements, customer-to-customer measurements, and network element fault monitoring, but there is a need to augment these functions.

There is no intent here to replace passive measurements, customer-to-customer measurements, and network element fault monitoring, but there is a need to augment these functions.

Probe Sequence Design

Probe sequence design sets the characteristics of both the distribution of packet sequences and the distribution of packets within a sequence. In particular, we define the size of packets, the sending rate for packets in each sequence, the time interval to send probe packet sequences, and the length of the test cycle (if different from the probe sequence length). Measurement design reflects a balance between the reliability of detecting events and the need to minimize network load during congestion and keep the measurements practical in server deployment.

The determination of the length of any test is governed by the events we wish to detect and characterize. Industry standards have long recognized that periods of degraded transmission lasting 10 s or more correlate with unavailability from the user's perspective (found in various standards, e.g., ITU-T G.821 and G.826). Furthermore, today's intelligent end terminals and network equipment will only wait in a degraded reception state for a limited time before taking some action (as described in ANSI T1.522-2000), and keep-alive timeouts on the order of seconds

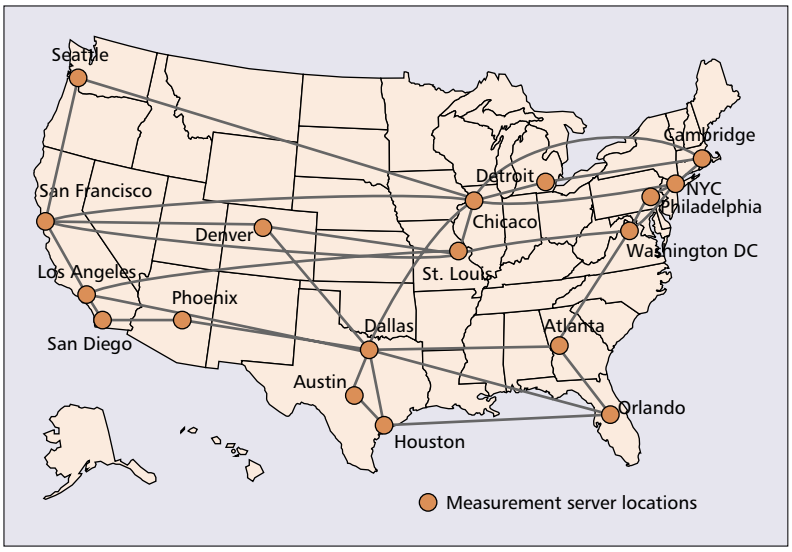

Figure 1. *Geographic positioning of measurement servers.*

or tens of seconds are sometimes taken as given. We make use of these precedents by assuming an event that causes significant degradation lasts at least 10 s.

Even if we assume that one degradation event occurs during every test cycle, there is a good chance that we will not detect it with a probe sequence whose duration is a small fraction of the test cycle (1 min in our case). A dense test that covers the whole test cycle would be ideal (providing continuous ability to characterize events with high accuracy), but practical considerations on the amount of test traffic preclude this. Thus, to at least detect events, our testing scheme includes a probe sequence at lower density to monitor time when the dense probe sequences are absent. A Poisson probe sequence running throughout the test cycle can provide near continuous detection. Thus, we have designed two test sequences, a periodic probe sequence and a Poisson probe sequence, consistent with RFC 3432 and RFC 2330, respectively. The Poisson probe serves to detect 10 s events with high confidence and is an unbiased measure of packet transport performance. The periodic probe sequence mimics a real-time VoIP application and can accurately characterize effects of repeating events on real-time application performance (nonrecurring events require less accurate assessment).

We divide each 24-hour day into 96 test cycles of 15 minutes. Each measurement server pair tests with a Poisson probe sequence of duration equal to the test cycle, and the following characteristics:

- Poisson distribution with average interarrival time of 3.3 s
- Packet size of 278 bytes, including headers (note that various packet sizes are present on the Internet)
- UDP protocol

Also, our measurement system launches two periodic probe sequences between pairs of measurement servers in every test cycle; each server initiates one sequence. The random start times are independent, so the sequences may overlap. The periodic sequences have the following characteristics:

- Interval of 20 ms between successive packets (or 50 packets/s)
- 1 min duration
- Random start time within the 15 min cycle
- Packet size of 60 bytes (including headers)
- UDP protocol

The probability of detecting and characterizing a single congestion event with the periodic probes is between 1/15 and 2/15, depending on whether they overlap (plus a small fraction due to the length of the congestion event, on the order of 0.01 for 10 s events).

The detection probability for the Poisson probe is simply the probability that at least one packet in the Poisson distributed sequence is sent (and lost) during the event. With all three probe sequences, the likelihood of detecting a congestion event is given by weighting the periodic and Poisson probe detection according to their overlaps in a conditional probability. The overall detection probability is slightly higher than the Poisson detection alone: 0.957 for 10 s degradation events and 0.522 for 2 s events.

Deployment in the AT&T U.S. Network

We have deployed this measurement system in the AT&T network using measurement servers placed in each of 18 major networking centers (Fig. 1 shows geographic locations). Each Poisson probe sequence operates between each of the 18 city pairs throughout the 15-minute cycle (18*17/2 = 153 tests). Each periodic probe sequence is randomly started from each of the 18 servers to the other 17 once every 15 min (306 tests every 15 min). This totals 29,376 periodic sequences each day. These measurements are being extended to the edge of the network (to the access routers).

The data are collected in a central server in the network care center (NCC). Raw data and summary data (containing 85 statistics) are collected for storage and analysis, and a high-level summary is processed for Web display.

Screen alerts based on loss, delay, and other measurements are presented to the NCC and the network operations center (NOC). Alarms (traps) based on this data (under conditions of thresholds and duration) are also sent to the NOC and NCC. The results display includes a high-level screen showing colored indicators at the city level with the capability to drill down to individual results and graphs, and enables efficient notification, diagnosis, and response.

Metrics

Along with a high-level summary, we collect a very detailed set of statistics from each test for more detailed analysis. Some of these metrics are traditional metrics such as mean round-trip delay, but others are new.

Delay

Delay is the time for packets to traverse the network from source to destination. The time interval includes serialization time, since it is calculated from first bit to last bit at the source and destination interfaces, respectively. We mea-

sure round-trip time to avoid introducing server time errors, but we also correct the round-trip measurement with processing time at the remote server. We collect several percentiles, the mean, minimum, and maximum delay for each test.

Loss

When a packet is sent, but does not arrive in a specified time, we designate that packet as lost. Our method waits 3 s before declaring the last packet lost, and a minimum of 3 s for all others up to the first (where we wait the entire length of the test plus 3 s). We can filter our results for cases where a constant waiting time is needed. We report the round-trip loss ratio for each test, as well as a count of degraded time intervals. In addition, we collect enough summary statistics from the raw data to determine if loss occurred in one direction, the extent of consecutive losses, and the loss pattern.

Traceroutes

Before each test a traceroute to the destination is performed. Tests are performed in both directions, so there is at least one traceroute in each direction every 15 min. The data is available by drilling down to the city-city specific data. Only traceroutes that have changed are kept to facilitate troubleshooting. The display thus shows all changed traceroutes for the 30 days prior to the date selected.

Delay Variation

There are two primary definitions of delay variation used in the industry and implemented in our system. The distribution of one-way delays with respect to a reference delay (e.g., the minimum delay of the population), $(R_n - T_n) - \min\{\delta_t\} = \delta t_n$, is the form of variation defined in ITU-T Y.1540, referred to here as *delta*, δ. On the other hand, the interpacket delay variation (IPDV) metric in RFCs 1889 and 3393 leads to the distribution of the differences in successive transfer times, or $(R_{n+1} - R_n) - (T_{n+1} - T_n) = \delta t_{n+1} - \delta t_n$. The distribution of this quantity is not equal to the distribution of the delay variation. If the assumption is made that the δt_n are independent and identically distributed variables, the variance of this distribution will be twice that of the desired distribution (the assumption of independence may overestimate the variance as there is likely to be correlation between the measurements due to congestion events). However, the IPDV metric is robust to reroutes during the measurement interval, while the one-way delta will report a larger variation because of the path delay change.

Although there are many advantages to using the strict delay variation (δ), it relies on accurate time-of-day synchronization of the server (and avoiding clock correction during the measurement). We thus use both metrics and collect them from our periodic probe sequences. Later we will show how these two metrics may differ.

Reordered or Out-of-Order Packets

Packet order is a property of successful packet transfer attempts, where the sending packet order is preserved on arrival at the destination. We defined a simple metric to determine if a network has maintained packet order, consistent with the IPPM framework in RFC 2330 and Y.1540 Appendix VII. The definition has two parts:

- Determine whether or not packet order is maintained.
- Quantify the extent of reordering.

The arrival order can be compared to the sending order through timestamps, message numbers, or byte stream payload numbers. The destination stores the "next expected" packet number based on previous packet arrivals or information exchanged at startup of the test. In-order packets have numbers greater than or equal to the "next expected" packet, and they set a new "next expected" value that cannot decrease (thereby requiring a nonreversing order). A reordered packet outcome occurs when the packet has a number lower than the "next expected." We note that packet losses do not cause reordering.

This metric classifies "late-arriving" packets as reordered. This is equivalent to the definitions in [2, 3].

We quantify reordering extent (part 2 of the definition) in multiple units of measure:

- Time
- Position
- Octets

When considering effects of reordering on applications, one should also use fundamental metrics (e.g., delay, delay variation, and loss). More detail may be found in the current work of the Internet Engineering Task Force (IETF) IP Performance Metrics Working Group [4].

Issues with the Current Measurement System

The system described above has several limitations, resulting from the many compromises made to ensure practical measurements. Some of these are:

- The lack of local high-accuracy time synchronization (e.g., GPS), so delay results are currently round-trip times (RTTs) only.
- Use of application-level measurements, not wire time or kernel time measurements
- 1 ms clock resolution
- Occasional measurement process interruptions
- Measurements within the ISP borders

Results

This section describes our measurement system's ability to detect, characterize, and help direct maintenance when necessary. A few detailed examples are shown to demonstrate the abilities and use of the system.

Low-Level Loss

With this measurement system we have been able to detect very low loss ratios, such as that caused by low bit error ratios on links or due to router card degradation (prior to failure). Passive measurements may also record this loss; however, the loss increases so gradually that this may not provoke action.

Before each test a traceroute to the destination is performed. Tests are performed in both directions, so there is at least one traceroute in each direction every 15 min. The data is available by drillling down to the city-city specific data.

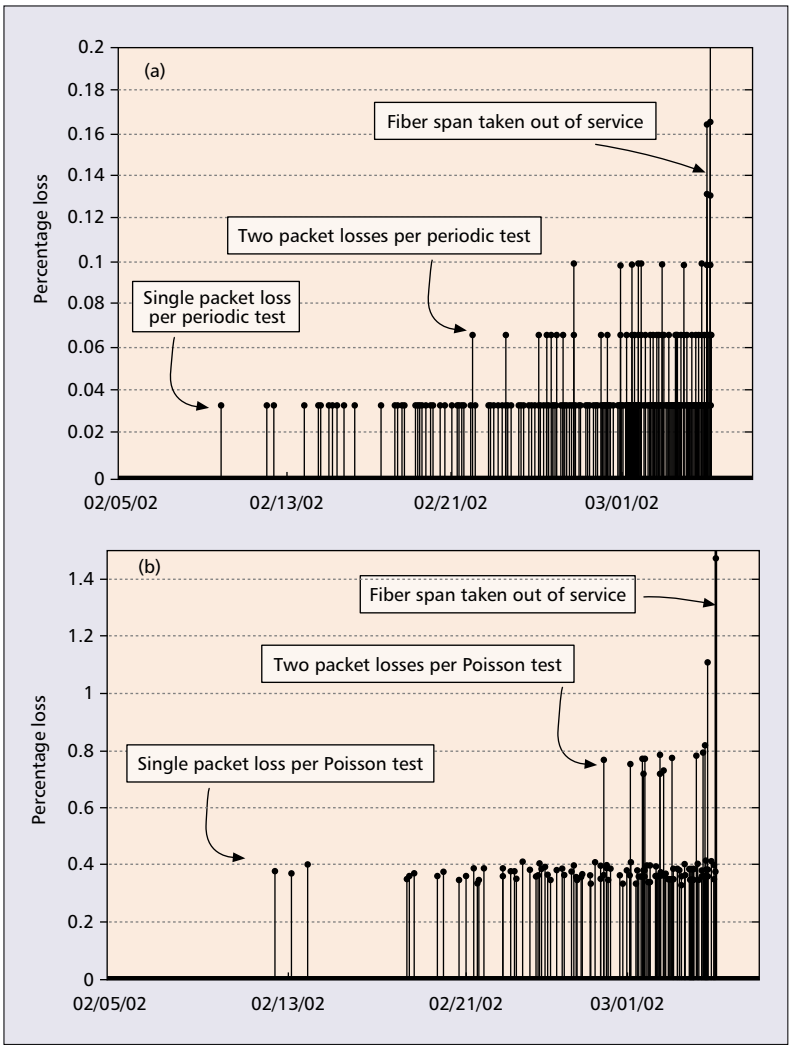

Figure 2. *Onset of low-level loss observed by a) periodic and b) Poisson probes.*

With the periodic probe sequence described here, it is possible to detect uniform degradation (e.g., a bit or card error) on the order of 0.03 percent loss ratio. The Poisson probe sequence, however, will only be able to detect uniform degradation of 0.3 percent loss. Thus, a small loss ratio measured consistently with the periodic sequence, and slightly more scattered tests with low loss with the Poisson sequence, is a symptom of low-level uniform loss.

Figure 2a and b show the percentage of loss (y-axis) per test as a function of time for a path with low-level bit errors on a single link. It can be seen that the onset of the problem was a uniform but very low probability of loss. The periodic probe recorded scattered low loss (single loss out of ~3000 packets), but as the loss probability approached 0.03 percent almost every test recorded the loss. The Poisson probe plot (b) shows a similar progression, except that only when the loss was near 0.37 percent did tests consistently record the loss (single packet). In this case the problem was identified before customers noticed any impact on their applications (there may have been some effect, especially to http and ftp, but probably not significant).

[1] *In the AT&T backbone these statistics are typically very close in value (within 1 ms, equal to our precision). This is due to the lack of congestion in the backbone. During congestion, queues lengthen and add as much as 500 ms delay. If a queue is full, the delay becomes more variable and is accompanied by loss.*

Route Changes

Another important type of network event that can be detected is a route change in the network. Route changes may occur because of outage (fiber cuts, equipment failure) or due to planned maintenance (card swaps, link cost changes). The active tests can determine the extent of loss during the event (layer 3 recovery) and the change in delay between the original and restored path. The AT&T network uses the Open Shortest Path First (OSPF) protocol for internal routing, so these results are particular to OSPF, although we believe that similar results will be true for other internal routing protocols such as IS-IS.

We have caught the reconvergence event with our tests several times, which has helped to estimate the effect of an OSPF reroute on a customer application stream. Typically, the Poisson sequence will be in progress during a reroute, and capture the resulting delay changes and packet loss. The periodic measurement reflects the path before, during, and after a reroute, but will not capture a loss burst unless it is running at the time of the incident. Below is a pictorial view of one such incident where the Poisson test captured the loss during the failure and reconvergence, while the periodic test coincided with the route returning to the normal path, in Fig. 3a. The idealized plot (a) shows statistics for the RTT delay per packet on the y-axis as a function of time (on the x-axis).

We see that the initial failure was the most damaging, causing five consecutive losses in the Poisson probe implying a burst of loss lasting about 15 s. Returning to the original route causes only 1 s of disruption, though, as characterized by the periodic probe. Subsequent analysis has shown us that this first reconvergence loss (due to failure of the current path) lasts from 5–30 s with the median time being 10–20 s. Layer 3 recovery is an attribute of IP routing design; this is a network recovering normally from a failure. It is only due to the continuous monitoring of path-level performance that we are able to detect and characterize these low-probability events. Other networks have also observed this behavior [5].

One characteristic of the reroute delay pattern is its "step-function" appearance. A very abrupt change in delay takes place, but except during the failure/rerouting, the minimum, mean, and 95th percentile of delay are very close together.[1] In Fig. 3b we show the minimum, mean, and 95th percentile of the RTT for each test (y-axis) as a function of time for the Poisson probe. We have confirmed that the traceroute for this path changed at the time of the incident.

Delay Variation

In this section we show the results of delay variation seen in the network. Most of the time the delay variation is undetectable (less than 2 ms IPDV range per test). However, a router interface timing misconfiguration led to variable delay through the router. As can be seen in Fig. 4, the average RTT changed only slightly. The IPDV range (maximum IPDV value in the test

minus the minimum IPDV value, line with diamonds) was on the order of 60 ms during the misconfiguration time. Also, the 99.9 percentile minus the minimum, or δ range, for each test (line with squares) corresponds to the shape of the IPDV curve, with a smaller but quite noticeable excursion from normal. Both of these delay variation metrics tracked the problem. Thus, the IPDV range captures the worst case delay per test, while the δ delay variation captures the deviation from the minimum. We have found that more common use of the jitter metric is when the jitter is not as dramatic as in this example, where only a few packets are disturbed that do not affect the 95th percentile or the mean drastically. Such data is useful to assess the impact of jitter on a real-time application stream.

Another use of the jitter metric is to estimate the buffer size needed to accommodate delay variations on individual packets. To produce the results in Fig. 5 we simulated traffic bursts in a laboratory test and view the result with different delay variation metrics. The burst traffic delayed a series of periodic packets, causing them to group in a burst while the queue empties. Figure 5 shows the RTT (solid line), δ values (line with solid triangles), and IPDV values (line with open circles) for the periodic packets. The first delayed packet has an IPDV value similar to the RTT and δ values (peak on graph), but the very next IPDV value is a packet that arrived *too early* by exactly the packet spacing (–20 ms). While the RTT and δ values decrement for successive packets (each is held 20 ms less than the previous), IPDV compares the original interpacket time to the closely spaced arrival times, and subsequent packets appear early. In cases where a single packet is held up by a number of milliseconds, the RTT and δ value will relax back immediately, but the IPDV value of at least two packets will be affected. Thus, the range of IPDV may overestimate the number of packets that would be lost in a fixed dejitter buffer. It appears that a dejitter buffer with 90 ms storage could accommodate this delay variation.

Forwarding Loops or Blenders

Although packet reordering is extremely infrequent on our network, the most intriguing event we have measured was first described in [1] where they named this phenomenon a *blender*. Blenders are caused by a transient routing loop that occurs when a router does not have the appropriate forwarding information and sends packets on a path that loops back instead of progressing toward the destination. New packets enter the loop until the router gets the routing update that enables it to forward the packets correctly. The result is a burst of reordered packets with varying RTTs. Casner observed an inverse relationship between IP time-to-live (TTL) values and transfer delay, where subsets of packets with the same (low) TTL had similar (high) delay results. While we observe similar RTT groupings, we do not collect the TTL for our probes. The blenders observed in Casner's experiment were between 4–14 s in duration, and exhibit both loss (possibly because of TTL expiration) and extensive reordering. While our data shows extensive reordering, we see less loss, and in general our events are shorter (~2–6 s). In Fig. 6 we show a blender characterized by our periodic probe sequence. In this figure the RTT in milliseconds (y-axis) is plotted in order of packet sending time (x-axis). The RTT values

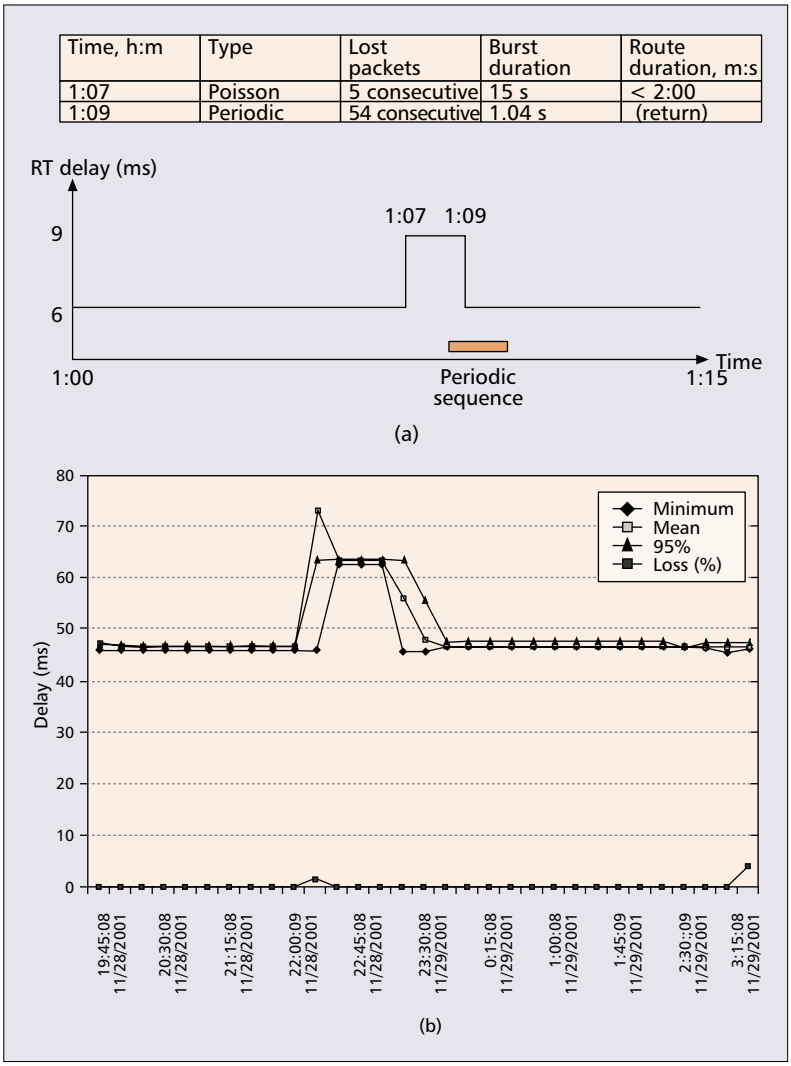

Figure 3. *Observations of loss during a route change: a) idealized; b) Poisson probe.*

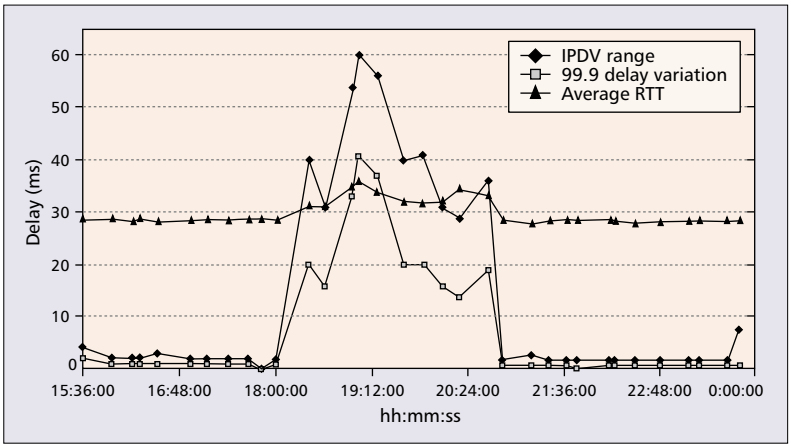

Figure 4. *Delay variation and IPDV range for timing-related jitter event.*

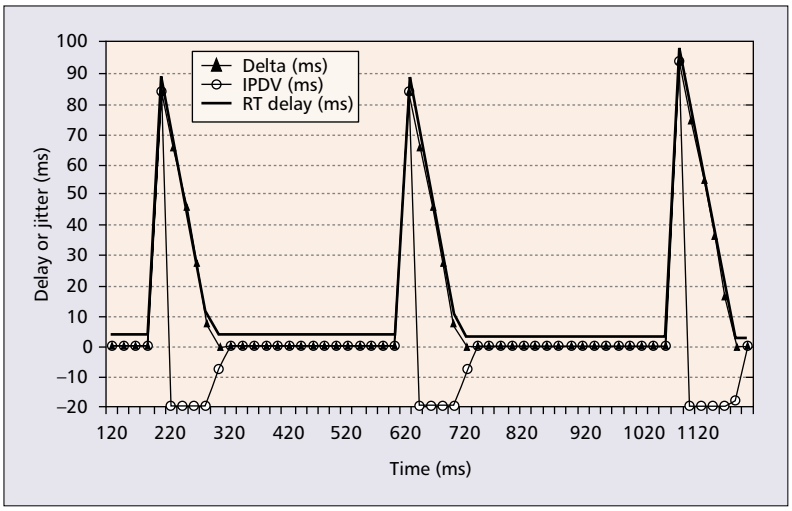

Figure 5. *Comparison of IPDV and δ delay variation (packet-by-packet measurements).*

appear in groups as a sequence of steps. The rise between steps is approximately uniform and reflects one trip around the loop. This indicates that the packet at the start of the event went through the loop path 22 times (number of steps), while later packets circled a smaller number of times. Lacking the TTL evidence we cannot confirm the hypothesis further; however, the circumstantial evidence and comparison with Casner's results leads us to believe that the above observation is a blender.

When we characterize this blender with the reordering metric described earlier, we find that it affected 88 packets: 79 were reordered, and 9 arrived in order (but had longer than usual delay). The maximum reordering extent was 85 packets, and maximum late time was 64 ms. Seven separate sequence discontinuities were observed, and no loss.

If we magnify the RTT of the initial packets sent into the loop, as shown in Fig. 6b, we observe a slope on the "tread" of the step. This feature is further support for a loop phenomenon, where no packets can leave, and the queue occupation grows over time at each hop in the loop. Another feature is that the width of the tread corresponds to the transit time of a single loop. There was a ~2 s interruption of packet flow while this blender loop was present, and this is the principal degradation real-time applications would experience.

Other Active Measurement Systems

There are several active measurement systems in place in the Internet. Very loosely, they can be organized into the following categories:

- Research systems whose aim is to characterize some aspects of IP performance (Internet2, Sprint, NIMI, Merit, etc.)
- Companies that provide a service to end customers by measuring the relative performance of ISPs, or Web hosting centers using their own system (e.g., Matrix, Keynote, Inverse)
- Systems based on an off-the-shelf hardware or software solution to an ISP or Web hosting center for monitoring performance (hardware: Brix, CQOS, RIPE; software: Agilent, CiscoWorks, Inverse)
- Internally developed single-network measurement systems such as the one described in this article (e.g., AT&T, Cable & Wireless, UUNET)
- Cross-provider measurements supported by multiple ISPs; the most prominent example: the RIPE NCC's Test Traffic Measurements Service (TTM) [6] in use in Europe and adjacent regions

All these systems have the ability to assess performance of a path, an advantage over passive measurement systems.

Some organizations have entries in multiple categories. Categories 2 and 3 have a strong commercial component, while 1 has less of a commercial component, and 4 and 5 have an operational motivation (although some data may be used for customer information purposes, e.g., at the AT&T site, http://www.att.com/ipnetwork, and other ISP sites).

Conclusions

In this article we describe a network-wide measurement system currently in operation on a tier 1 ISP backbone. We show through examples how this system contributes to network management, complementing traditional element-level monitoring. Our active measurement system also keeps detailed data for later analysis. In other words, this is a permanent measurement infrastructure with historical data.

We describe our measurement design, with the goal of detecting and characterizing network performance degradation events in dimensions relevant to nontraditional IP network applications. Our sampling design strikes a balance between statistical accuracy and collection times similar to typical user session length.

The platform allows quick deployment and evaluation of new metrics for packet transfer performance in addition to the standard metrics. Recent additions to the system include the reordering metrics, various jitter metrics, and loss pattern metrics, including the degraded second/minute metric.

The value of this detailed data is shown with a sample of events, including the effect of low-level loss, the effect of route changes, and the rare event of route loops. These events were easily identified and examined using data collected by our system.

We have also learned some important general rules from continuously monitoring the network. These may be used for network design, management of operational procedures, or vendor management. Continuous evaluation of network performance leads us to the opinion that while periods of degradation are undesirable, they often can be coped with, whereas instability (such as is seen during route changes) can be more damaging, especially to real-time applications.

We have found that active monitoring of a network provides valuable proactive alerts of

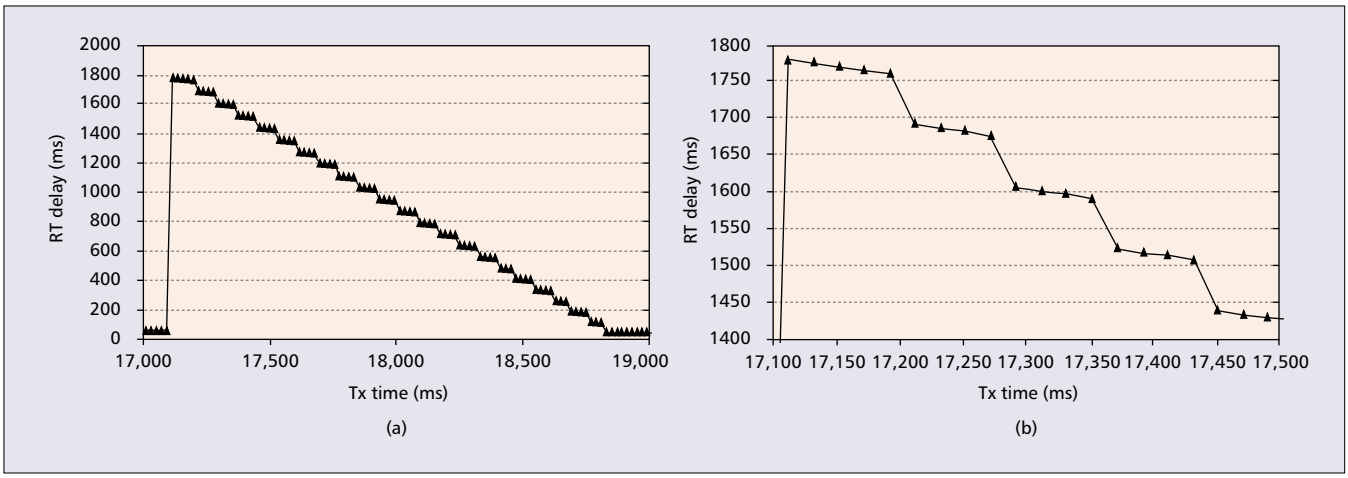

Figure 6. *a) A simple blender; b) a magnified view of blender steps.*

impending degradation. It can serve as a useful complement to element-level (passive) monitoring if engineered well. Such a system provides useful information for network design and maintenance. We have shown that large-scale networks can deploy appropriate active measurements effectively in support of operations.

Acknowledgments

We would like to acknowledge the people who helped make this version of the measurement system a reality: Ron Kulper, George Holubec, Shashi Pulakurti, Mai-Uyen Nguyen, Nicole Kowalski, Arvind Ramarajan, Mark Perkins, and Ganga Maguluri. We also thank Bill Halloran, Chuck Dvorak, and Luis Morales for their continued support.

References

Note that the ITU-T Recommendations and IETF RFCs cited in this article are available from their Web sites.

[1] S. Casner, C. Alaettinoglu, and C. Kuan, "A Fine-Grained View of High Performance Networking," NANOG 22 Conf., May 20–22, 2001; http://www.nanog.org/mtg-0105/agenda.html

[2] D. Loguinov and H. Radha, "Measurement Study of Low-bitrate Internet Video Streaming," *Proc. ACM SIGCOMM Internet Measurement Wksp. 2001*, San Francisco, CA, Nov. 1–2, 2001.

[3] V. Paxson, "Measurements and Analysis of End-to-End Internet Dynamics," Ph.D. dissertation, UC Berkeley, 1997; ftp://ftp.ee.lbl.gov/papers/vp-thesis/dis.ps.gz

[4] Morton *et al.*, "Packet Reordering Metric for IPPM," IETF, work in progress.

[5] G. Iannaccone *et al.*, "Analysis of Link Failures in a Large IP Backbone," IMW 2002.

[6] http://www.ripe.net/ttm/

Additional Reading

[1] V. Jacobson, "Congestion Avoidance and Control," *Proc. Sigcomm '88 Symp.*, Aug. 1998.

[2] M. Mathis *et al.*, "The Macroscopic Behavior of TCP Congestion Avoidance Algorithm," *Comp. Commun. Rev.*, vol. 27, no. 3, July 1997.

[3] Raisanen, Grotefeld, and Morton, "Network Performance Measurement with Periodic Streams," RFC 3432, Nov 2002.

[4] V. Paxson *et al.*, "Framework for IP Performance Metrics," RFC 2330, May 1998.

Biographies

LEONARD CIAVATTONE (lencia@att.com) is a senior engineer at AT&T Laboratories responsible for performance and characterization testing of AT&T IP network elements. He has specialized in the prototyping and certification of QoS functionality for both data and voice IP services. He has developed numerous software test tools, including a large-scale active probe sequence measurement system and monitor for the AT&T ISP network. Prior to AT&T he worked at Lucent Technologies on interoperability and performance evaluation testing of multivendor QoS-based IP and ATM switching environments. He began his data networking career in 1986 when he joined ECI Telecom (formerly Telematics International) where he specialized in performance characterization testing and source code support of frame relay and X.25 switching software.

ALFRED C. MORTON [M] (acmorton@att.com) is a technology consultant at AT&T Laboratories, the research and development engine of AT&T. He has been a recognized contributor and editor in U.S. and international network performance standards committees for over 20 years. He currently serves as Co-Chair of the Benchmarking Working Group of the IETF. He also participates in the work of IETF's IP Performance Metrics Working Group, ITU-T Study Groups 12 and 13 on multimedia and IP performance, and Technical Subcommittee T1A1 (Performance, Reliability, and Signal Processing). After working with Computer Sciences Corp. and the U.S. Army Satellite Communications Agency, where he was a project team leader responsible for test and evaluation of satellite and terminal systems, he joined AT&T Bell Laboratories in 1984. While at Bell Laboratories he worked on facsimile transmission, data application performance analysis, customer opinion modeling, videoconferencing quality measurements, and network timing distribution and standards, and was appointed a Distinguished Member of Technical Staff in 1995. He earned his M.S. in electronic engineering at Monmouth College in 1983, after receiving his B.S.E.E. there in 1977. His thesis topic was an interference rejection system for direct-sequence spread-spectrum communications. He holds two U.S. patents.

GOMATHI RAMACHANDRAN (gomathi@att.com) is a principal technical staff member at AT&T Laboratories. Her major area of interest is network performance assessment, particularly the design of measurements and systems to manage, track, and analyze performance. Prior to her career at AT&T, she did her postdoctoral work in modeling DNA dynamics at the Courant Institute in New York, and before that received a Ph.D. in physical chemistry from Cornell University, Ithaca, New York, for work on quasi-classical ion-molecule dynamics. She holds a Master's degree in chemistry from the Indian Institute of Technology (IIT), Bombay, and a Bachelor of Science degree from Bangalore University, India.