Measuring the net

Markus Peuhkuri

2003-10-16

Lecture topics

- Why network is measured
- How network can be measured
- What is measured
- How one can utilize measurements
- *IP* networks assumed
- Focus on *quality-related* measurements, no discussion about security-related monitoring such as IDS systems.

Who cares about measurements in network [6]

- ISP
  - capacity planning
  - operations
  - value add services (e.g. customer reports)
  - usage-based billing
  - equipment and network performance evaluation

  ⋆ bandwidth utilisation
  ⋆ packets per second
  ⋆ round trip time (RTT)
  ⋆ RTT variance
  ⋆ packet loss
  ⋆ reachability
  ⋆ circuit performance
  ⋆ routing diagnosis
- Users: corporations and individuals
  - monitor performance
  - plan upgrades
  - negotiate service contracts
  - set user expectations
  - optimise content delivery
  - usage policing

  ⋆ bandwidth availability
  ⋆ response time
  ⋆ packet loss

    - ⋆ reachability
    - ⋆ connection rates
    - ⋆ service qualities
    - ⋆ host performance

- Vendors
  - improve design and configuration of equipment
  - implement real-time debugging and diagnosis of deployed hardware
    - ⋆ trace samples
    - ⋆ log analysis

Measurements provide insights relating to [15]

- Network provisioning
- Peering arrangements
- Per-customer accounting and SLA verification
- Per-per accounting (traffic balance of trade)
- Performance management
- Tracking topology and routing changes
- Tracing DoS attacks
- ATM/cell/packet/circuit level errors and other troubleshooting
- Connectivity complexity and vulnerability
- TCP flow dynamics
- Routing table and address space efficiency

Operator requirements for measurements

- Network is a long-time invertment
- Operations must have continuity
- Need for common standards to collect measurement data. For example, it is not sufficent just to have common protocol transfer measurements but also data collection must be uniform: any inconsistencies in statistical definitions, protocol levels, or data collection should be avoided. For example, are layer 2 headers and framing or IP headers included in byte counts?
- Measurement system must scale as network grows and transmission rates increases
  $\Rightarrow$ Data must be aggregated as much as possible
- Measurements must not interfere with data transmission

Network operator time scales

The demand for measurements depends on the timescale it is used for

Months network planning, network extension or introducing new technologies to meet future needs for capacity and reliability

Hours or days capacity management: the network is reconfigured to optimise utilisation

Real-time apply short-term corrections to network configuration in event of congestion or failure automatically or manual

Network metrics categories [12]

Utilisation metrics: packet and byte counts, peak metrics, protocol, and application distribution.

Performance metrics: round-trip time (at different layers) and packet drop count.

Availability metrics: long-term line, route or application availability.

Stability metrics: short-term fluctuations that degrade performance such as line status transitions, route changes, next hop stability and short term ICMP anomalous behaviour.

Measurement types

Active measurements: Test traffic is sent

- data is sent, either real application data or measurement-only data
- transfer time (or possible data loss) is measured
  - in both ends, needs syncronised clocks
  - on sending end the response (round-trip-time)
- adds traffic to network
- does the test traffic have different treating?

Passive measurements: Existing traffic is used

- existing traffic is captured
- adds no extra traffic to network (expluding possible result transfer)
- some route cannot be measured if there is no traffic

Both techniques can be combined

Active measurements

- A data is sent to network (addressed to some host)
- Other system (not necessary the destinated host) may
  1. timestamp
  2. reply
- Sender records reply (possibly)
- Standard tools, or
- Special soft- and/or hardware
  Examples of measurement platforms are:
- NLANR AMP `http://watt.nlanr.net/`
- DREN AMP `http://www.sd.wareonearth.com/amp` (AMP peer network)
- Internet End-to-End Performance Monitoring at SLAC `http://www-iepm.slac.stanford.edu/`
- National Internet Measurement Infrastructure `http://www.ncne.nlanr.net/nimi/`
- RIPE's Test Traffic Measurements `http://www.ripe.net/test-traffic/index.html`
- Surveyor `http://www.advanced.org/surveyor/`
- CAIDA s Skitter Project `http://www.caida.org/Tools/Skitter/`

Active measurement tools

ping uses ICMP echo request/echo response

- a host sends ICMP echo requst, other system replies with echo response
- round-trip time and packet loss

+ each IP host *must* implement ICMP echo server
  ⇒ no need to additional software
- *but*, many firewalled hosts are broken, furthermore in many cases it is possible to learn that system is on network even if it does not reply to ICMP messages
- systems implement limit of ICMP messages sent per second to protect for Denial-of-Service attacks ⇒a missing reply may not be because of network loss
- ICMP processing may be in lower priority task

UNIX simple services echo, discard, chargen

- diagnostics tools for TCP and UDP
- often disabled or rate-limited as can be used for DoS

Traceroute finds out forward path

- sends UDP, TCP or ICMP datagrams with increasing TTL, starting from TTL=1
- a router possibly[1] sends ICMP time exceeded message pack if TTL goes to zero
  ⇒ each datagram travels one router further

HTTP-request measures application performance

- a document is requested from a web server and time needed to transfer is measured
- the server may have considerable effect: the server may be heavily loaded or there may be delays in connections to backend servers (databases etc.) if page is dynamicly created.
- other services may be used also

IP Performance Metrics (ippm) [16]

- IETF working group developing a set of standard metrics for Internet data delivery services
  - quality
  - performance
  - reliability
- Can be used by all parties: network operators, end users, or independent testing groups
- Metrics defined:
  - connectivity [13]
  - one-way delay and loss [1, 2]
  - round-trip delay and loss [3]
  - delay variation [9]
  - loss patterns [11]
  - packet reordering
  - bulk transport capacity [14, 18]
  - link bandwidth capacity

The IPPM WG will develop a set of standard metrics that can be applied to the quality, performance, and reliability of Internet data delivery services. These metrics will be designed such that they can be performed by network operators, end users, or independent testing groups. It is important that the metrics not represent a value judgement (i.e. define “good” and “bad”), but rather provide unbiased quantitative measures of performance.

[1] See discussion about ping above

Problems with active measurements

- Different level of service for different protocols. For example, the web traffic (port 80) may have higher priority than network news (port 119).
- Some types of traffic may be administratively blocked by firewall systems: this results a false negative in connectivity tests. Also some types of traffic may have some kind rate limit.
- Application traffic profile may be different from test traffic: the application fidelity may not be easily derived from simple loss and delay figures but one must know also *which ones* are lost. For example, a 5 % packet loss may result severe frame loss (more than 50 %) for video traffic [7].
- Periodic stream test traffic, bursty application traffic. At times of high load, when there can be QoS problems and large amount of application traffic is carried, the proportion of test packets is low. This results in underestimating the times of low QoS [10].

Passive measurements

- Network traffic is directed to measurement device
  - shared medium, for example non-switched Ethernet
  - optical/electrical signal divided by splitter/tap. Optical splitter directs a part of signal in fibre (ratios 50/50 – 90/10, attentuation $\approx$ 4/4 – 1/12 dB) to another fibre. These are sensitive to wavelength. Electrical taps have some amplifing circuit.
  - pass-through device receives data and retransmits it. This introduces additional point of failure.
  - port mirroring in router or in switch: traffic is copied to monitoring port.
- Traffic is captured from network
  - full census
  - random sampling
  - deterministic sampling
- Data is recorded for post-processing or analysed real-time
  - per-packet analysis. For example, protocol and packet length distribution, packet interarrival times.
  - per-flow analysis. Traffic is grouped into flows (see below) and statistics are collected for each flow.

What is a flow

- A flow is a series of packets travelling from one part of network to another part of network

  unidirectional A→B different from B→A

  bi-directional A→B same as B→A

  It is not (always) possible to observe both directions at the same location, because of asymmetric routing.
- Potential granularities [17, p. 60]
  - application, identified by
    - TCP or UDP port numbers
    - transport protocol
    - IPSec SPI [4]
    - IPv6 flow identifier
  - host, identified by
    - network layer address (IP address)

    * link layer address (e.g. MAC address)
    * hostname (e.g. DNS name)
  - network, identified by
    * address prefix
    * AS number
    * domain name
    * arbitrary group of hosts
  - traffic sharing a common path in the network, identified by
    * link (interface on router)
    * ATM or FR virtual channel identifier
    * MPLS path
    * AS path

- The most common granularities
  - (source address, source port, protocol, destination address, destination port)
  - (source network, destination network)
  - (destination network), this is how a routing takes place!
- Packets belonging to the same flow *should* receive similar performance, especially if granularity is high.
  - variying performance is bad for many protocols and applications

Flow lifetime

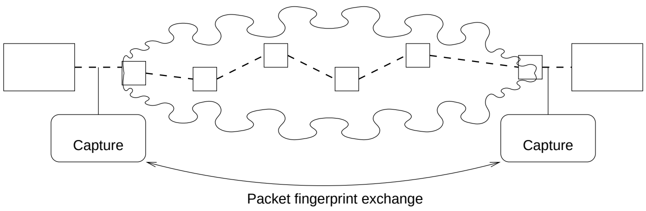
- Lifetimes vary
  - two packets exchanged in few milliseconds: one DNS query
  - millions of packets in a month: several TCP connections between two servers
- Flow timeout depends on application, more on following lectures

Problems in passive measurements

- Sensitive data may be exposed. Legistlation varies by country, but in general the operator is allowed to learn about traffic for maintaining network and troubleshoting problems, but it should be very limited. In Finland, *unauthorized* wiretapping may result in a fine or even up to three years of imprisonment.
  - user data sensitive: passwords, credit card numers, and other confidential info – use of encrypted connections (ssh, https) help for this.
  - IP addresses sensitive. An IP address can identify user to a person or household – what "content" are you using?
  - other protocol fields possibly sensitive. For example, in a small network simply a list of protocols used (TCP and UDP port numbers) can be sensitive.
- Data volume can be huge: If one considers operator backbone link that may be currently STM-16 (OC-48, 2.5 Gbit/s) and it may have on average packet size of 512 bytes. In a full load situation there will be more than 600,000 packets per second. If 64-byte record is stored for each packet (timestamp and header) that would result 36 MiB/s data stream and for the next faster network, STM-64 (OC-192) or 10GE this figure is four times larger. Compare this to measurements from traditional circuit-switched networks where 200 bytes are more than enough to save essential information from telephone call. If average call is 3 minutes in duration, the record is only 0.01 % of data flow.
  - sampling can reduce data a lot: some metrics seem to survive sampling very well, especially for count-based random sampling. See IETF PSAMP working group for more info.
- Misbehaving end systems. One must be careful when analysing traffic. Not all implementations work as specifications indicate but there is a large number of errors in applications and operating systems.

Multi-point measurements

- Provide additional information about network traffic
  - routing
  - per-packet delays
- Information exchange and mapping additional problem
  - fingerprinting
  - trajectory sampling



Non-network measurements

- Network application logs
  - http servers
  - mail servers
  - ftp servers
- Response time for application, for example to monitor database server; includes both network and application delays. These can be used as part of SLA verification tools, especially if "whole service" (i.e. both the network and the server) is provided by one service provider.
- VoIP phone quality. VoIP calls are graded either by humans or some automated system to evaluate perceived quality of VoIP connections. There are many methods for this, starting from simple S/N tests (not very good to evaluate packet-based voice) to standardized/proposed ones like PSQM, PSQM+, and PESQ. [8]

Service Level Agreement Measurements

- Network performance has a value
  ⇒ need to verify that one is getting quality paid for
- Large range of SLA definitions
  - the service availability verification "is accomplished by the Operator pinging the Customer's router"
  - threshold values for
    - available bandwidth
    - packet loss rate
    - packet delay

    contribute for the service level that is
    - satisfactory
    - degraded
    - unavailable

    If you promise 99.9999 % availability, you *must* also *define availability*

    Availibility and maximum downtime

%	per year	per day
99	3d 15h 36m 0s	14m 24s
99,9	8h 45m 36s	1m 26s
99,99	52m 33s	8.6s
99,999	5m 15s	0.9s
99,9999	32s	0.1s

Although "six nines" may be feasible for single high-availability network device, for a large network or a long network path it is very hard requirement.

- SLA measurement systems
  - set of soft- or hardware agents around network
  - do tests at times, typicaly retreives some web pages few times a hour, similary performing DNS queries or ICMP Echos.
  - report results for server
  - user can retrive reports and receive alerts
  - not yet according to IPPM

Accounting

- Operator may have volume- and class-based charging
- Needs to now how much each customer has traffic
- Possibly different price for different targets: which portion of traffic by a customer is local, domestic, global, or served by cache systems.
- Commonly done using cflowd on routers
- IETF IPFIX working on standard flow information exchange (not yet any RFCs, only I-Ds)

Which measurement strategy to select

Who you are?

Tier-1 operator use of special hardware feasible. As backbone operator may have only hundred or so nodes, even if single device is expensive (to measure high-speed links), one needs only small number of devices.

Tier-2 operator special hardware on selected links. Tier-2 operator provides services for tier-3 operators and very large customers.

Tier-3 operator only partial coverage for measurements. Tier-3 operator, especially one focused to small business and consumers, may have hundreads of thousands of links to watch. There is no any change to install special hardware even to small fraction of links.

Corporate user monitors its own usage and checks for received quality. For a corporation, it is important that network is available and provides sufficient service so that network does not became limiting factor of business. It may be important also to monitor network so that usage is along guidelines.

Home users does not have knowledge to measures. For an average user, it is very hard to identify what is the problem if "web is broken". It may be problem on network or at user computer. There is a need for easy-to-use tools for non-professionals to identify problems.

Summary

- Measurements provide information about network

  active: what kind of service additional traffic would receive

  passive what the *present-day* traffic looks like, what kind of service it receives

- It is important to select proper measurement
- . . . and to interept readings right
- If you plan to provide QoS, you *must* measure

For a large list of tools, see [5].

References

[1] G. Almes, S. Kalidindi, and M. Zekauskas. A One-way Delay Metric for IPPM. Request for Comments RFC 2679, Internet Engineering Task Force, September 1999. (Internet Proposed Standard). URL:`http://www.ietf.org/rfc/rfc2679.txt`.

[2] G. Almes, S. Kalidindi, and M. Zekauskas. A One-way Packet Loss Metric for IPPM. Request for Comments RFC 2680, Internet Engineering Task Force, September 1999. (Internet Proposed Standard). URL:`http://www.ietf.org/rfc/rfc2680.txt`.

[3] G. Almes, S. Kalidindi, and M. Zekauskas. A Round-trip Delay Metric for IPPM. Request for Comments RFC 2681, Internet Engineering Task Force, September 1999. (Internet Proposed Standard). URL:`http://www.ietf.org/rfc/rfc2681.txt`.

[4] L. Berger and T. O'Malley. RSVP Extensions for IPSEC Data Flows. Request for Comments RFC 2207, Internet Engineering Task Force, September 1997. (Internet Proposed Standard). URL:`http://www.ietf.org/rfc/rfc2207.txt`.

[5] Internet tools taxonomy. On web, `http://www.caida.org/tools/taxonomy/`. Refered 2002-05-05. URL:`http://www.caida.org/tools/taxonomy/`.

[6] K. Claffy and T. Monk. What's next for internet data analysis? status and challenges facing the community. *Proceedings of the IEEE*, 85(10):1563–1571, October 1997.

[7] Michele Clark and Kevin Jeffay. Application-level measurements of performance on the vBNS. In *ICMCS, Vol. 2*, pages 362–366, 1999.

[8] Adrian E. Conway and Yali Zhu. A simulation-based methodology and tool for automating the modeling and analysis of voice-over-IP perceptual quality. *Performance Evaluation 54 (2003) 129 147*, 54(2):129–147, October 2003.

[9] C. Demichelis and P. Chimento. IP Packet Delay Variation Metric for IP Performance Metrics (IPPM). Request for Comments RFC 3393, Internet Engineering Task Force, November 2002. (Internet Proposed Standard). URL:`http://www.ietf.org/rfc/rfc3393.txt`.

[10] Jorma Jormakka ja Kari Heikkinen. QoS/GOS parameter definitions and measurement in IP/ATM networks. In *Proceedings of First COST 263 International Workshop, QofIS 2000*, pages 182–193, Berlin, Germany, September 2002.

[11] R. Koodli and R. Ravikanth. One-way Loss Pattern Sample Metrics. Request for Comments RFC 3357, Internet Engineering Task Force, August 2002. (Informational). URL:`http://www.ietf.org/rfc/rfc3357.txt`.

[12] M. Lambert. A Model for Common Operational Statistics. Request for Comments RFC 1857, Internet Engineering Task Force, October 1995. (Informational) (Obsoletes RFC1404). URL:`http://www.ietf.org/rfc/rfc1857.txt`.

[13] J. Mahdavi and V. Paxson. IPPM Metrics for Measuring Connectivity. Request for Comments RFC 2678, Internet Engineering Task Force, September 1999. (Internet Proposed Standard) (Obsoletes RFC2498). URL:`http://www.ietf.org/rfc/rfc2678.txt`.

[14] M. Mathis and M. Allman. A Framework for Defining Empirical Bulk Transfer Capacity Metrics. Request for Comments RFC 3148, Internet Engineering Task Force, July 2001. (Informational). URL:`http://www.ietf.org/rfc/rfc3148.txt`.

[15] Preliminary measurement spec for internet routers. Draft at `http://www.caida.org/tools/measurement/measurementspec/`. Work in process.

[16] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis. Framework for IP Performance Metrics. Request for Comments RFC 2330, Internet Engineering Task Force, May 1998. (Informational). URL:`http://www.ietf.org/rfc/rfc2330.txt`.

[17] Markus Peuhkuri. Internet traffic measurements – aims, methodology, and discoveries. Licentiate thesis, Helsinki University of Technology, Finland, May 2002.

[18] V. Raisanen, G. Grotefeld, and A. Morton. Network performance measurement with periodic streams. Request for Comments RFC 3432, Internet Engineering Task Force, November 2002. (Internet Proposed Standard). URL:`http://www.ietf.org/rfc/rfc3432.txt`.