

Traffic Measurement and Analysis (1)

SOI ASIA Lecture 2002/11/26

Kenjiro Cho
Sony Computer Science Labs, Inc.
kjc@csl.sony.co.jp

1

measurement goals

- for operations
 - trouble shooting
 - diagnosis and tuning of performance, reliability
 - usage report
 - long-term planning of capacity, equipment, cost evaluation
- for protocol/software/hardware engineering
 - trade-off in design (e.g., buffer size vs. cost)
 - to verify things are working as designed
 - to look for unexpected (important in Internet)
- for scientific interests (new discoveries)
 - characteristics of delay, throughput, loss
 - modeling (e.g., TCP, web traffic)
 - self-similarity/fractal traffic
 - ▷ abundant data, simulation tools

2

measurement needs combined skills

- goals could be operational, engineering, scientific
 - all unseparable, all skills required
 - ▷ knowledge of operational environment
 - ▷ engineering of measurement tools
- output can be facts, findings, new ideas
 - new ideas are not always necessary
 - facts, especially long-term measurement, are valuable
- but you should have clear goals
 - better to start with real problems to solve
 - ▷ there are many issues and problems but some are more important than others

3

why traffic measurement of Internet is so hard?

- massive, diverse and changing traffic
- mechanisms at different layers in different time scale
 - interact with each other
- dynamics
 - Internet mechanisms are adaptive and resilient
 - traditional measurement techniques are often not applicable
- pathological traffic is not unusual
 - by bugs, misconfigurations, errors, mismatches, accidents
- we still don't have good understanding

4

massive volume of traffic

- unprecedented scale with unprecedented growth
 - e.g., traffic volume: 100Mbps traffic
 - ▷ 12MB/sec 715MB/minute 42GB/hour 1TB/day
- far more data than we can analyze
 - techniques needed to reduce data size
 - ▷ filtering: e.g., record only TCP SYN packets
 - ▷ aggregation: e.g., flow-based accounting
 - ▷ sampling: e.g., record 1 in n packets
- still, details matter
 - a big impact often comes
 - ▷ from small fraction
 - ▷ from minor differences

5

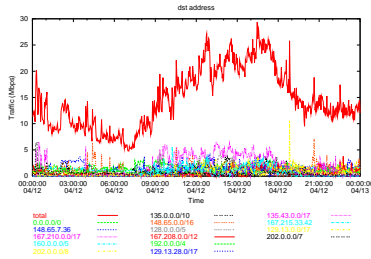
diverse traffic

- large variation in traffic mix between sites
- backbone vs. access links
 - access line types: fiber, ADSL, modem, wireless, satellite
 - ▷ differences in bandwidth, delay, loss
- typical traffic doesn't exist!

6

constant change of traffic pattern

- daily, weekly traffic pattern
- trend changes over time
 - web completely changed traffic pattern
- hard to predict future!



7

time scale of traffic management

- long-term
 - capacity planning
- day
 - pricing (off-time rate discount)
- session
 - service pricing, admission control, routing
- round-trip time
 - end-to-end flow control, various timeout mechanisms
- packet time
 - packet scheduling
- less than packet time
 - link layer dependent

8

Internet dynamics

- dynamics
 - packet switching
 - statistical multiplexing
 - queueing
 - feedback mechanisms
 - at different layers in different time scale
 - e.g., TCP congestion control
- scaling property
 - Internet traffic is bursty
 - correlation, long-range dependences
 - traditional measurement techniques often not applicable
 - e.g., independent (memoryless) events, random sampling
 - median and 90th-percentile more useful than mean and stddev
 - try log-scale plots to see scaling property

9

other issues

- problems often occur at boundaries of different networks
 - cooperation needed but not easy
- operators vs. researchers
 - different interests and culture
 - build good relationship
- cost: measurement doesn't come free
 - willingness to invest
- privacy in traffic data
- companies often do not publish results

10

commonly-used management tools

- quick overview
 - ping
 - reachability, round-trip time
 - traceroute
 - path detection
 - tcpdump
 - packet capturing
 - SNMP
 - usage monitoring

11

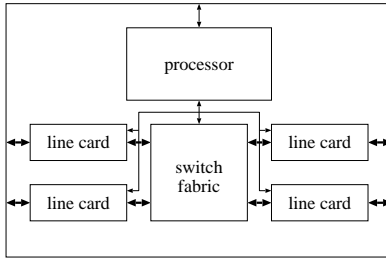
ping

- a popular and widely-available tool to check connectivity
- ICMP-echo request/reply
- limitations
 - ping responses do not mean network is working correctly
 - ICMP is not representative of host/network performance

12

router architecture

- fast path: hardware assisted processing
- slow path: software processing



13

ping sample output

```
% ping -c10 www.ait.ac.th
PING www.ait.ac.th (202.183.214.46): 56 data bytes
64 bytes from 202.183.214.46: icmp_seq=0 ttl=113 time=220.550 ms
64 bytes from 202.183.214.46: icmp_seq=1 ttl=113 time=241.832 ms
64 bytes from 202.183.214.46: icmp_seq=2 ttl=113 time=228.779 ms
64 bytes from 202.183.214.46: icmp_seq=3 ttl=113 time=220.574 ms
64 bytes from 202.183.214.46: icmp_seq=4 ttl=113 time=219.312 ms
64 bytes from 202.183.214.46: icmp_seq=5 ttl=113 time=217.608 ms
64 bytes from 202.183.214.46: icmp_seq=6 ttl=113 time=218.355 ms
64 bytes from 202.183.214.46: icmp_seq=7 ttl=113 time=221.564 ms
64 bytes from 202.183.214.46: icmp_seq=8 ttl=113 time=218.330 ms
64 bytes from 202.183.214.46: icmp_seq=9 ttl=113 time=219.085 ms
```

```
--- www.ait.ac.th ping statistics ---
10 packets transmitted, 10 packets received, 0% packet loss
round-trip min/avg/max/stddev = 217.608/222.599/241.832/7.084 ms
```

14

traceroute

- exploit TTL (time-to-live) of IP
 - router returns ICMP TIME EXCEEDED to the sender when TTL becomes 0
- limitations
 - path may change over time
 - path may be asymmetric
 - reports one of the interfaces of router

15

traceroute sample output

```
% traceroute www.ait.ac.th
traceroute to www.ait.ac.th (202.183.214.46), 64 hops max, 44 byte packets
 1 entry (133.138.1.2) 0.350 ms 0.308 ms 0.297 ms
 2 foundry2.otemachi.wide.ad.jp (133.138.0.1) 0.961 ms 1.603 ms 1.553 ms
 3 cisco5.otemachi.wide.ad.jp (203.178.140.220) 181.694 ms 203.383 ms 199.252 ms
 4 210.132.94.77 (210.132.94.77) 1.807 ms 1.953 ms 1.713 ms
 5 gsr-ote1.kddnet.ad.jp (203.181.96.37) 1.872 ms 1.663 ms 2.350 ms
 6 tr-ote109.kddnet.ad.jp (203.181.96.74) 2.362 ms 2.198 ms 2.147 ms
 7 210.132.93.186 (210.132.93.186) 214.827 ms 218.536 ms 215.359 ms
 8 202.47.253.145 (202.47.253.145) 218.988 ms 217.795 ms 216.383 ms
 9 202.47.252.190 (202.47.252.190) 216.720 ms 217.435 ms 217.882 ms
10 202.183.160.121 (202.183.160.121) 216.964 ms 216.934 ms 216.781 ms
11 www.ait.ac.th (202.183.214.46) 219.197 ms 229.315 ms 217.640 ms
```

16

tcpdump

- packet capture tool
 - capture the first N bytes of packets
- flexible filtering
 - e.g., capture only TCP SYN from host X
- enables detailed analysis
- limitations
 - huge volume
 - difficult to capture high-speed links

17

tcpdump sample output

```
14:53:24.878901 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: S 1758089118:1758089118(0) win 57344
  cmsg 1440,nop,wscale 0,nop,nop,timestamp 77899233 0>
14:53:24.882544 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: S 959813947:959813947(0) ack 1758089119 win 16384
  cmsg 33160,nop,wscale 0,nop,nop,timestamp 0 77899233>
14:53:24.882597 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: ack 1 win 58548 <nop,nop,timestamp 77899233 0>
14:53:24.904230 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: P 1:451(450) ack 1 win 58548 <nop,nop,timestamp 77899236 0>
14:53:24.910081 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: P 1:324(323) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.912819 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 324:1764(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.912846 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: ack 1764 win 56785 <nop,nop,timestamp 77899236 0>
14:53:24.919719 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 1764:3204(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.920049 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 3204:4644(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.920888 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: ack 4644 win 53905 <nop,nop,timestamp 77899237 0>
14:53:24.921624 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: ack 4644 win 58001 <nop,nop,timestamp 77899237 0>
14:53:24.922177 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 4644:6084(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.929047 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 6084:7524(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.929086 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: ack 7524 win 55668 <nop,nop,timestamp 77899238 0>
14:53:24.930272 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 7524:8964(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.931643 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 8964:10404(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.931692 linux.cs1.sony.co.jp.4804 > www.wide.ad.jp.http: ack 10404 win 52788 <nop,nop,timestamp 77899238 0>
14:53:24.936876 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 10404:11844(1440) ack 451 win 16384 [flowlabel 0xc0414]
14:53:24.938104 www.wide.ad.jp.http > linux.cs1.sony.co.jp.4804: 11844:13284(1440) ack 451 win 16384 [flowlabel 0xc0414]
```

18

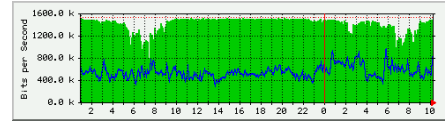
SNMP (Simple Network Management Protocol)

- SNMP allows a remote user to
 - query information, store information, set traps
 - by UDP (unreliable)
- standardized set of traffic statistics
 - supported by most of routers, switches, host OS
 - many management/monitoring products
- MIB (Management Information Base)
 - tree structured database of SNMP objects
 - e.g., interfaces.ifTable.ifEntry.ifOutOctets
 - standard MIBs and private MIBs
 - get, set, get-next to access MIB
- limitations
 - supported statistics are limited
 - most counter statistics are hard-coded: e.g., interface counters
 - accessing to MIB objects is expensive

19

MRTG

- popular tool to show SNMP data
- time series data aggregated over time
 - daily, weekly, monthly
- inbound/outbound traffic
 - can be used for other types of time series data
- RRDtool: successor of MRTG



20

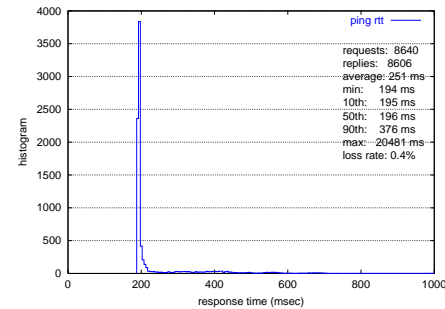
real-world data

- one day ping measurement
 - from home in Tokyo to university in East Coast
- facts
 - 8640 queries, 8606 replies, 0.4% loss
 - average rtt: 251ms
 - stddev: 391ms
 - min rtt: 194ms
 - 10th: 195ms
 - 50th: 196ms
 - 90th: 376ms
 - max: 20481ms

21

ping round-trip time histogram

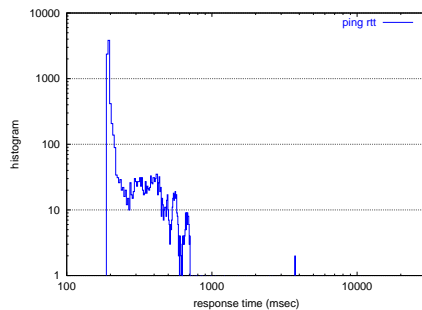
- most replies: 194-196ms



22

ping round-trip time histogram in log scale

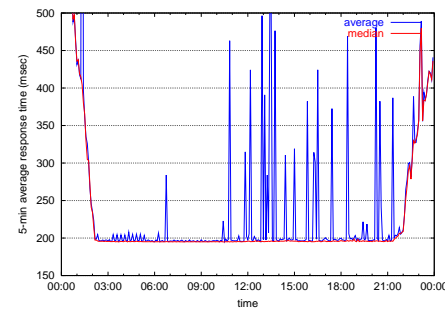
- log-plot to see scaling property
 - no scaling property in data



23

daily plot (mean and median)

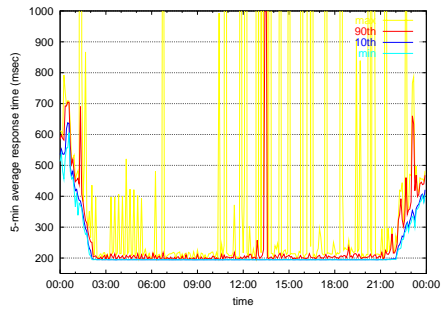
- median more stable than mean
- service degrades during 22:00 - 26:00



24

tails of distribution

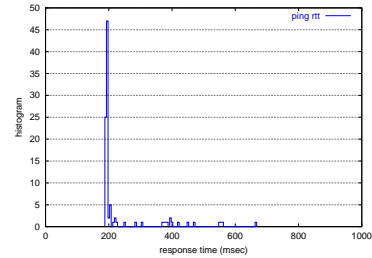
- daily plot (min, 10th, 90th, max)
 - minimum, maximum
 - 10th-percentile, 90th-percentile to remove outliers



25

limitations of histogram

- needs appropriate bin size
 - too small: each bin doesn't have enough samples (e.g., empty bins)
 - too large: only few regions available
- enough samples needed
 - histogram with 100 samples



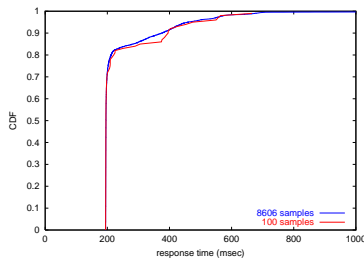
26

cumulative distribution function (CDF)

- density function: probability of observing x

$$f(x) = P[X = x]$$
- cumulative distribution function: probability of observing x or less

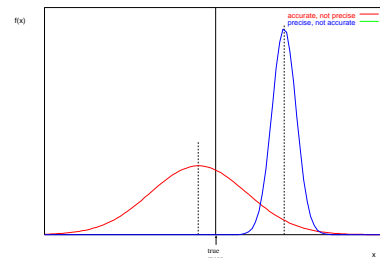
$$F(x) = P[X \leq x]$$
- better than histogram when
 - sample count is not enough or outliers are not negligible



27

sample average revisited

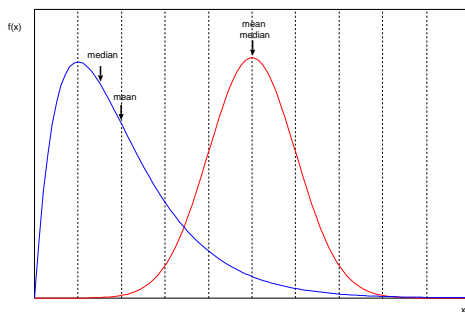
- accuracy
 - how close to true (population) mean
- precision
 - variance in data



28

mean and median

- not equal if distribution is asymmetric



29

average (mean)

average over n sample values

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

variance

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

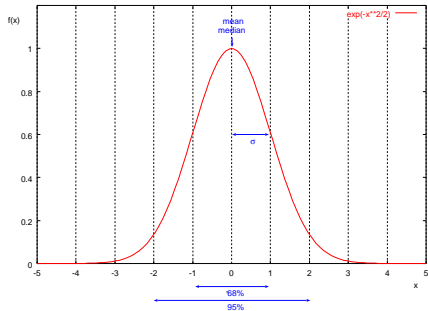
standard deviation s (in the same unit as mean)

- sample mean (for large samples) follows normal distribution
 - central limit theorem (see statistics textbook)

30

normal distribution

- 68% within (mean-stddev, mean+stddev)
- 95% within (mean-2*stddev, mean+2*stddev)



31

confidence interval

- confidence interval for the mean
 - provides probabilistic bounds
 - tells how much uncertainty in the estimate

$$Prob\{c_1 \leq \mu \leq c_2\} = 1 - \alpha$$

(c1, c2): confidence interval
 100(1 -): confidence level

- e.g., with 95% confidence, the population mean is between c1 and c2
 - traditionally, 95 or 99% is used for confidence level

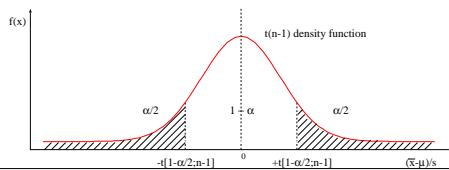
32

confidence interval (cont'd)

- from central limit theorem
 - if observations are independent and samples come from the same population with mean μ and standard deviation
 - then, sample mean for large samples is normal distribution with mean μ and standard deviation σ/\sqrt{n}

$$\bar{x} \sim N(\mu, \sigma/\sqrt{n})$$

- increase sample size to get more accuracy
- $(\bar{x} - \mu)/(s/\sqrt{n})$ for samples from normal populations
 - follows t(n-1) distribution



33

how to use confidence interval for mean

- applications
 - provide confidence interval to show possible range of mean
 - from sample mean and stddev, compute how many trials are needed
 - to satisfy a given confidence interval
 - repeat measurement until a given confidence interval is reached
- summary:
 - be careful when you use average
 - sometimes, average is not so useful in Internet measurement

34

measurement techniques

- management tools are useful but not designed for measurement
- next lecture:
 - types of measurement
 - throughput, delay, path, routing
 - data reduction techniques
 - filtering, aggregation, sampling
 - clock and timestamp

35

summary

- overview of measurement issues
 - operational, engineering, scientific skills are needed
- popular management tools
 - ping, traceroute, tcpdump, SNMP
- using real ping data
 - histogram, CDF
- mean and confidence interval

36