MEASUREMENT: MONITORS & A FIRST LOOK AT INTERNET MEASUREMENTS

Gaia Maselli

maselli@di.uniroma1.it

SAPIENZA
Università di Roma

Performance measurement

To measure the performance of a system you need at least two tools

- A tool to load the system: *load generator*
- A tool to measure the results: *monitor*

SAPIENZA
Università di Roma

Monitor

- A monitor is a tool used to observed the activities on a system
- Monitors observe the performance of systems, collect performance statistics, analyze the data, and display results
- Reasons to monitor a system
- To optimize system performance
- To measure resource utilization
- To find the performance bottleneck
- To tune the system
- To characterize the workload

Terminolgy

- **Event**: a change in the system state (ex. Arrival of a packet)
- **Trace**: a log of events usually including the time of the event, the type of the event, and other important parameters associated with the event
- **Overhead**: consumption of system resources to run the monitor (ex. CPU or storage)
- **Domain**: the set of activities observable by the monitor
- **Input rate**: the maximum frequency of events that a monitor can correctly observe
  - Burst mode: specifies the rate at which an event can occur for a short duration
  - Sustained mode: the rate that the monitor can tolerate for long durations
- **Resolution**: the coarseness of the information observed (ex. A monitor may be able to record time only in units of 16 milliseconds)

Monitor classification

- **Event driven**: is activated only by the occurrence of certain events
- **Timer driven** (**sampling** monitor): is activated at fixed time intervals by clock interrupts
1. **Software** monitors
   - Issues in buffer size (they record data in buffers and then into disk or other storage), data compression and analysis (save memory but adds overhead)
2. **Hardware** monitors: a separate piece of equipment attached to the system being monitored (no system resources are consumed)
3. **Firmware** and hybrid monitors (network monitoring where existing network interfaces can be easily microprogrammed to monitor all traffic on the network)

SAPIENZA
UNIVERSITÀ DI ROMA

Monitor classification (cont)

Distributed-system monitors: several components that work separately and concurrently

Ex. To determine the link with the highest error rate, the errors at each and every link in a network should be observed.

- *Observation*: raw data gathering
  - **Implicit spying**: promiscuously observing the activity of the system (packet sniffer)
  - **Explicit instrumenting**: incorporating trace points, probe points, hooks, or counters in the system
  - **Probing**: making request on the system to sense its current performance
- *Collection*: data gathering (from a group of observer)
  - Synchronization issues: time stamp from different observer cannot be compared unless the observers' clocks are close to each other within some tolerance
- *Analysis*: can be performed in the observer or the analyzer depending on the specific case(highest error rate => analyzer)

Measurement

Performance evaluation technique which can be applicable to any existing network or prototype

- Internet
- Sensor network (terrestrial & underwater)
- RFID systems
- …

Internet measurement

Internet measurement

Motivation

- Many quantitative measures of the internet are absent
  - Internet is not the result of a centralized design
  - It is constantly changing in size, configuration, traffic, and applications mix
- How big is the Internet?
- How much traffic flows over the Internet?
- What is the structure of the Internet?
- What are the statistical properties of network traffic?
- What demands do different applications place on the network?
- What is the capacity of the path to my server?

Measurement issues

- Internet devices do not always provide the kind of measurements that are most useful for understanding the network
- Collecting measurements of the Internet can result in *huge datasets* that are difficult to store, transfer, process, and analyze
- Commercial service providers often *do not s*hare information about the internal details of their networks
- No possibility for remote monitoring
- No knowledge of what is needed
- Some forms of internet measurement can *violate privacy* and *raise security concerns*
- *Synchronization* is an important issue

SAPIENZA
Università di Roma

Where can measurements be made?

- At every point of the Internet network

- Inside LAN

- In and around an Internet Service Provider (organization controlling one or multiple autonomous systems)

    - At backbone routers, access routers, gateway routers, peering routers

- At network access points (NAP) or exchange points of multiple ISP

SAPIENZA
UNIVERSITÀ DI ROMA

Inside LAN

- Carried out for local test-beds
- Typically not of significant interest in the "Internet" measurement sense
- Measurements of local latency and hardware related measurements are done on a LAN
- Security reasons

Inside a backbone

- ISPs constantly monitor the network for a variety of purposes
  - Ensuring availability
  - Scanning for outages or attacks
  - Topology changes
  - Compliance with service level agreements
  - Traffic trends (diurnal, weekday, and other periodicities)
- Measurements inside the backbone are done from an *intra-organizational point of view*
  - Help in proper provisioning of resources
  - Indicate whether router and link upgrades may be warranted
- Tools: SNMP and packet tracing

Key objectives of backbone measurements

- Capacity planning
  - To see if additional points of presence (PoPs) are needed
  - To see if existing PoPs in a network need additional capacity
  - Requires packet loss, delay and throughput measurements

- Provisioning (proper allocation of bandwidth to various links inside an ISP)
  - Different applications have varying sensitivity to latency (consider email vs. multimedia streaming)
  - Typically ISP over-provision the network, but the appropriate degree of over-provisioning has to be computed
  - Requires measurements of changes in traffic patterns at small scale and traffic across all pairs of PoPs

SAPIENZA
Università di Roma

Key objectives of backbone measurements (cont)

- Link utilization
  - Long term information on link utilization help to identify PoP that need more capacity
  - Understanding the traffic associated to a set of costumers, an organization can provide tailored service to them
  - Requires packet delay measurement with high speed monitors
- Proper tuning of interior gateway protocol (balance traffic)
  - Monitor links and routing information to get a view of eventual traffic shifts (sudden growth in a particular protocol)
  - Requires combining traffic matrix and routing data

SAPIENZA
Università di Roma

Key objectives of backbone measurements (cont)

- Security
  - Attacks and anomalies
  - Requires monitoring significant changes in traffic patterns between PoPs
  - Requires examine link utilizations periodically and notice significant increases
- Identify failures
  - Failures that are not related to attacks may affect availability
  - Path failures trigger rerouting

SAPIENZA
Università di Roma

Entry points into a network: *gateway* routers

- ## Access control
  - Entry points of the network are the first line of defense and the best place to filter out unwanted traffic

- ## Overall statistics
  - The *netflow* tool allows to export *per-flow summaries* of traffic which includes information about
    - start and ending time of flows,
    - duration,
    - source and destination IP and ports
    - autonomous systems
    - Fraction of traffic destined to customers inside the network
    - Portion of traffic that is transiting through the network

SAPIENZA
Università di Roma

Entry points into a network: *peering* router

- Monitoring inter-domain connectivity
- Ensure balanced traffic exchange
  - Traffic volumes exchanged between private peers has be approximately equal
  - Deviations from expectation can trigger policy decisions (peering at other points) or require the exceeding peer to pay for the surplus traffic
- Monitoring BGP
  - Examining convergence
  - Fixing problems
  - Locating routing loops
  - Faults can be deliberately injected to examine their impact, such as how long before the route is repaired, a better path is discovered, etc.

Entry points into a network: *access* router

- Access router connect the backbone to the set of customers
- Access routers are also the routers used to connect to web and mail server
- Availability is crucial (failure rate must be low or non-existent)
- Some customers may require packet filtering
- Some customers may require periodic statistics and constant performance monitoring to ensure that anomalous events are kept to a minimum
- Many customers expect the access provider to look for attacks and sudden fluctuations in network traffic ina pro-active manner

Entry points into a network: *exchange points*

- An Internet Exchange Point permits various ISPs to exchange traffic

- There are commercial, government and research/ education exchange points

- One of the primary purposes of a network exchange point is to keep traffic local, i.e., move traffic between two participants without having to route it through long distance routes

- Measuring at exchange points allows to get a broader idea of shifts in traffic patterns (increasing online gaming, etc.)

Wide area network

- The measurement sites examined so far are restricted to a single location or point of presence
- The amount of traffic at each of these places may differ significantly
- Any internet topology related measurement has to carried out in the wide area
- Measurements on a wide area network: across the Internet on multiple locations
  - Coordinated and carried out simultaneously
  - Separately over a period of time

WAN: various places in the network

- WAN measurements are done by researchers and measurement companies
- Typical locations: all the ones listed in the previous slides but across a wider area: at multiple point of presence on the Internet

WAN: multi-site measurements

- A collection of nodes are used for simultaneous and cooperative measurement
- Example: to obtain a measure of how typical users might experience a Web site, measurements might be carried out on several locations on the Internet corresponding to different user populations
- Carrying out multi-site measurements in a coordinated fashion may require
  - clock synchronization
  - Execution serialization
  - A command and control mechanism capable of handling access and resource control
- Available platforms: NIMI and PlanetLab
- Representativeness: user populations, choices of clients, servers, etc. are reasonably well represented so that proper inferences can be made

SAPIENZA
UNIVERSITÀ DI ROMA

Role of time

- *Synchronization* is the process of ensuring that physically distributed processor have a common notion of time
- Many measurement tasks require accurate time measurement
    - Packet round trip time
    - Packet delay across routers and over links
    - Producing a time-ordered view of measurements taken at different places in the network
    - Response time and throughput
- Synchronization is a challenging issue as the Internet is a distributed system with components often separated by considerable distances
- In the case an accurate clock exists, the distance between component induces communication latency which can make clock readings stale by the time they arrive

SAPIENZA
UNIVERSITÀ DI ROMA

Definition

- t: **true time** at any instant
- $C(t)$: **apparent time** reported by a clock at time t
- *Offset* of a clock $\theta(t)=\theta(t)-C(t)$: the difference between the time it reports and the true time.

  An accurate clock has always $\theta(t)=0$

- *Rate* of a *clock $\gamma(t) = dC(t)/dt$* : the first derivative of its apparent time with respect to true time.

  An accurate clock has $\gamma(t)$ close to *1*

- *Skew* = $\gamma-1$ is the difference between its rate and the correct rate

SAPIENZA
UNIVERSITÀ DI ROMA

Observations

- Accuracy is a more stringent requirement than zero skew
- A clock that has large offset (is inaccurate) but has **zero skew** is still useful for certain type of measurement
  - Packet round trip time
  - Packet inter arrival time
- Measurement with one-away packet delay or time ordering of events occurring in different places reuqires clocks with **zero offset**

Sources of time information

External time sources

• Radio services that disseminate time information

1. Radio clocks

2. Global Positioning System (GPS): a constellation of 32 satellites in 12-hour orbit, available worldwide

   • Requires large antennas

   • outdoor

3. CDMA cellular phone system

   • Indoor and outdoor

   • Available in areas having CDMA telephony providers (not in Europe)

Sources of time information

PC-based clocks

• Standard PCs have two clocks

1. A battery-powered *hardware clock* keeps track of time when the system is turned off, and is not typically used when the system is running

2. *Software clock* is the usual source of time while the system is running

   • To read it: gettimeofday() or GetsystemTime()

3. Time Stamp Counter (TSC) register which is incremented every processor cycle

Synchronized time

- Many Internet measurement tasks involve measurements taken using different clocks.

- Ex. One way packet delay measurements involve measuring the departure time of a packet in one location and the arrival time of the packet at another point.

- Accurate determination of the true delay can be obtained in one of two ways:

1. Using synchronized clocks => Network Time Protocol (NTP)

2. Inferring clock offsets and removing their effects after measurements are made (see Paxson ACM SIGMETRICS 98)

SAPIENZA
UNIVERSITÀ DI ROMA