The “Internet”: Challenges

- Explosive growth and moving target
- Heterogeneity any which way you look
- New requirements: security, reliability, ...
- Complex user behavior and traffic dynamics
- Designed as open, cooperating, static system
- Highly interacting systems
  - Temporal: between users, hosts and networks
  - Spatial: among different components
  - Vertical: across different networking layers
**Scenarios To Be Addressed**

**Example 1:**
- Situation: network connectivity fails
- Presumed action: call system administrator
- Effect: no phone call possible
- Why: telephone service via VoIP

**Example 2:**
- Situation: remote computer/router fails
- Presumed action: connect to device via control net.
- Effect: no connection
- Why: no reliable backup network

---

**The “Network”: Advantages**

- Highly engineered structure
  - Well specified and documented
- Unique measurement capabilities
  - In theory unlimited access to data ☺
- Exploiting available data
  - Learn from the existing artifact
  - Use invariants and structural models not details
  - Consider emerging phenomena
  - Use network wide data
Challenges

The total is more than the sum of its pieces

- Specify and manage **services** rather than **components**

- Address the gap in understanding between **individual pieces** and the overall

Internet: A Complex Layered Distributed System

- **Physical connectivity**: Links
- **Point-to-point connectivity**: NIC, switches
  - distributed hardware, protocols – local management
- **End-to-end connectivity**: Routers
  - Forwarding, addressing, routing
  - Distributed hardware, protocols, software, management by Internet Service Providers (ISPs)
- **Process-to-process connectivity**: TCP, UDP
  - De-/multiplexing, reliability, congestion control, ...
- **Applications**: Web, P2P, ...
  - Users
  - Distributed, independent, autonomous, ...
Tools

- Instrumentation and analysis
  - Integrate measurements into the design process
  - Collect data at a variety of different locations/levels
  - Find invariants and correlate various datasets

- Simulation
  - Build a mirror world for “what if” studies
  - Verify explanations

- Test-Labs
  - Incorporate variability
  - Provide an friendly/unfriendly environment

Measurement Analysis

- Interesting when combining multiple datasets
  - Same data different locations
    - Root cause analysis of BGP events
    - Internet topology derivation
    - ...
  - Different data sets
    - Intra-domain traffic matrix
    - Inter-domain traffic matrix
    - ...
Terminology And General Issues

- Measurements vs. metrics
- Measurement capabilities
- Collection of measurement data
- Data reduction techniques
- Clock issues

Terminology: Measurements vs. Metrics

- active measurements
  - topology, configuration, routing, SNMP
  - packet and flow measurements, SNMP/RMON
- end-to-end performance
  - average download time of a web page
  - TCP bulk throughput
  - end-to-end delay and loss
  - link bit error rate
  - link utilization
  - traffic matrix
  - demand matrix
- state
  - active topology
  - active routes
- traffic
Terminology And General Issues

- Measurements vs. metrics
- Measurement capabilities
- Collection of measurement data
- Data reduction techniques
- Clock issues

Active Measurements

- Definition:
  - Injecting measurement traffic into the network
  - Computing metrics on the received traffic
- Scope
  - Closest to end-user experience
  - Least tightly coupled with infrastructure
  - Comes first in the detection/diagnosis/correction loop
Passive Measurements

- **Definition:**
  - Observing traffic into the network
  - Computing metrics on the monitored traffic

- **Scope**
  - Closest to network
  - Tightly coupled with infrastructure

---

**Passive Measurement Capabilities:**

**Packet Monitors**

- **Available data:**
  - All protocol information
  - All content
Monitoring A LAN Link

- Shared media (Ethernet, wireless)
  - Host A
  - Host B
  - Monitor

- Multicast switch
  - Host A
  - Switch
  - Host C
  - Monitor

- Monitor integrated with a bridge
  - Host A
  - Bridge/Monitor
  - Host B

Monitoring A WAN Link

- Splitting a point-to-point link
  - Router A
  - Router B
  - Monitor

- Line card that does packet sampling
  - Router A
Passive Measurement Capabilities: Packet Monitors (2.)

- Available data:
  - All protocol information
  - All content
- Possible analysis:
  - Application performance
  - User behavior (search engine comparisons)
  - Application usage (P2P usage)
  - Abuse detection (intrusion detection system)
- Disadvantages:
  - Data flood
  - Data aggregation
  - Needle in a haystack
  - Only captures on network information (no device info)
  - Usually needs fixed installations

Selecting Traffic

- Filter to focus on a subset of the packets
  - IP addresses/prefixes (e.g., to/from specific Web sites, client machines, DNS servers, mail servers)
  - Protocol (e.g., TCP, UDP, or ICMP)
  - Port numbers (e.g., HTTP, DNS, BGP, Napster)
- Collect first n bytes of packet (snap length)
  - Medium access control header (if present)
  - IP header (typically 20 bytes)
  - IP+UDP header (typically 28 bytes)
  - IP+TCP header (typically 40 bytes)
  - Application-layer message (entire packet)
IP Header Format

<table>
<thead>
<tr>
<th>4-bit Version</th>
<th>4-bit Header Length</th>
<th>8-bit Type of Service (TOS)</th>
<th>16-bit Total Length (Bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-bit Identification</td>
<td>3-bit Flags</td>
<td>13-bit Fragment Offset</td>
<td></td>
</tr>
</tbody>
</table>

| 8-bit Time to Live (TTL) | 8-bit Protocol | 16-bit Header Checksum |

| 32-bit Source IP Address |

| 32-bit Destination IP Address |

Options (if any)

Payload

Analysis Of IP Header Traces

- Source/destination addresses for traffic
  - Identity of popular Web servers & heavy customers
- Traffic breakdown by protocol (TCP/UDP/ICMP)
  - Amount of traffic not using congestion control
- Distribution of packet delay through the router
  - Identification of typical delays and anomalies
- Distribution of packet sizes
  - Workload models for routers and measurement devices
- Burstiness of the traffic on the link over time
  - Provisioning rules for allocating link capacity
- Throughput between each pair of src/dest addresses
  - Detection and diagnosis of performance problems
TCP header format

- 16-bit source port number
- 16-bit destination port number
- 32-bit sequence number
- 32-bit acknowledgement number
- 16-bit TCP checksum
- Options (if any)
- Payload
- 16-bit source port number
- 16-bit destination port number
- 32-bit sequence number
- 32-bit acknowledgement number
- 16-bit TCP checksum
- Options (if any)
- Payload

Tcpdump Output
(three-way TCP handshake and HTTP request message)

617756405:617756405(0) win 32120 <mss 1460,sackOK,timestamp 46339
0,nop,wscale 0> (DF)

2598794605:2598794605(0) win 16384 <mss 512>

1:1(0) ack 1 win 32120 (DF)

1:513(512) ack 1 win 32256 (DF)

513:676(163) ack 1 win 32256 (DF)

1:179(178) ack 676 win 16384
TCP Header Analysis

- Source and destination port numbers
  - Popular applications (HTTP, Napster, SMTP, DNS)
  - Number of parallel connections between source-dest pairs
- Sequence/ACK numbers and packet timestamps
  - Out-of-order/lost packets; violations of congestion control
  - Estimates of throughput and delay of Web downloads
- Number of packets/bytes per connection
  - Size of typical Web transfers; frequency of bulk transfers
- SYN flags from client machines
  - Unsuccessful connection requests; denial-of-service attacks
- FIN/RST flags from client machines
  - Frequency of Web transfers aborted by clients

Packet Contents

- Application-layer header
  - HTTP and RTSP request and response headers
  - FTP, NNTP, and SMTP commands and replies
  - DNS queries and responses; OSPF/BGP messages
- Application-layer body
  - HTTP resources (or checksums of the contents)
  - User keystrokes in Telnet/Rlogin sessions
- Security/privacy
  - Significant risk of violating user privacy
  - More sensitive for information from higher-level protocols
  - Traffic analysis thwarted by use of end-to-end encryption
HTTP Request And Response Message

Request
GET /tutorial.html HTTP/1.1
Date: Mon, 27 Aug 2001 08:09:01 GMT
From: jrex@research.att.com
User-Agent: Mozilla/4.03
CRLF

Response
HTTP/1.1 200 OK
Date: Thu, 12 Jul 2001 10:09:03 GMT
Server: Netscape-Enterprise/3.5.1
Last-Modified: Sun, 12 Mar 2000 11:12:23 GMT
Content-Length: 23
CRLF
Traffic measurement talk

Application-layer Analysis

- URLs from HTTP request messages
  - Popular resources/sites; potential benefits of caching
- Meta-data in HTTP request/response messages
  - Content type, cacheability, change frequency, etc.
  - Browsers, protocol versions, protocol features, etc.
- Contents of DNS messages
  - Common queries, frequency of errors, query latency
- Contents of Telnet/Rlogin sessions
  - Intrusion detection (break-ins, stepping stones)
- Routing protocol messages
  - Workload for routers; detection of routing anomalies
  - Tracking the current topology/routes in the backbone
Mechanics: Application-level Messages

- Application-level transfer may span multiple packets
  - Demultiplex packets into separate “flows”
  - Key of source/dest IP addresses, port #s, and protocol
  - Hash table to store packets from different flows

![Diagram showing packet demultiplexing process]

- Reconstructing ordered, reliable byte stream
  - Sequence number and segment length in TCP header
  - Heap to store packets in correct order & discard duplicates

- Extraction of application-level messages
  - Parsing the syntax of the application-level header
  - Identifying the start of the next message (if any)

- Logging or online analysis of message
  - Record URL, header, body, checksum, timestamps, etc.
  - Copy traces or analysis result to separate machine
Placement Of The Monitor (Edge)

- **Client edge**
  - Capture all traffic to/from a single group of clients
  - Useful for evaluating effectiveness of a proxy
  - May not be representative of other clients

- **Server edge**
  - Capture all traffic to/from a set of Web sites
  - Useful for detailed characterization of access patterns
  - May not be representative of accesses to other sites

Placement Of The Monitor (Core)

- **Middle of network**
  - Capture all traffic traversing a particular link
  - Useful for capturing a diverse mix of traffic
  - May not see all traffic traveling from host A to host B
  - May not see the reverse traffic from host B to host A
System Constraints

- High data rate
  - Bandwidth limits on CPU, I/O, memory, and disk/tape
  - Could monitor lower-speed links (near the edge of network)
- High data volume
  - Space limitations in main memory and on disk/tape
  - Could do online analysis to sample, filter, & aggregate
- High processing load
  - CPU/memory limits for extracting, counting, & analyzing
  - Could do offline processing for time-consuming analysis
- General solutions to system constraints
  - Sub-select the traffic (addresses/ports, first \( n \) bytes)
  - Kernel and interface card support for measurement
  - Efficient/robust software and hardware for the monitor

Passive Measurement Capabilities:
Packet Monitors (3.)

- Deployment scenarios:
  - Needs cooperation of the network operator
  - Limited number
  - Specialized hardware/software
  - Data collection / aggregation infrastructure
- Challenges
  - Data integrity
  - Incomplete data
  - User privacy & network security
  - Data correlation
  - Data privacy vs. data sharing
  - Data filtering
  - Data collection across network confederations
Passive Measurement Capabilities: Flow Statistics

- Available data:
  - Summary information about traffic flows

IP Flows: What Is It?

- Set of packets that “belong together”
  - Source/destination IP addresses and port numbers
  - Same protocol, ToS bits, ...
  - Same input/output interfaces at a router (if known)

- Packets that are “close” together in time
  - Maximum spacing between packets (e.g., 15 sec, 30 sec)
  - Example: flows 2 and 4 are different flows due to time
Passive measurement capabilities:
Flow statistics (2.)

- Available data:
  - Summary information about traffic flows
- Possible analysis:
  - (Application performance)
  - User behavior
  - Application usage (P2P usage)
  - Abuse detection (intrusion detection system)
- Disadvantages:
  - Coarser grain information
  - Data flood
  - Data aggregation
  - Needle in a haystack
  - Only captures on network information (no device info)
  - Usually needs to be configured on network devices

Passive Measurement Capabilities:
Flow Statistics (3.)

- Deployment scenarios:
  - Needs cooperation of the network operator
  - Larger number
  - Specialized hardware/software
  - Data collection/aggregation infrastructure
- Challenges
  - Lack of detail
  - Data integrity
  - Incomplete Data
  - Data correlation
  - Data privacy vs. data sharing
  - Data collection across network confederations
Passive Measurement Capabilities: SNMP/RMON statistics

- Available data:
  - Summary information from and about devices

SNMP/RMON

- Definition:
  - Standardized by IETF
  - SNMP = Simple Network Management Protocol
  - Definition of management information base (MIB)
  - Protocol for network management system to query and effect MIB

- Scope:
  - MIB-II: aggregate traffic statistics, state information
  - RMON1 (Remote MONitoring):
    - More local intelligence in agent
    - Agent monitors entire shared LAN
    - Very flexible, but complexity precludes use with high-speed links
SNMP: Naming Hierarchy And Protocol

- Information model: MIB tree
  - Naming & semantic convention between management station and agent (router)

- Protocol to access MIB
  - Get, set, get-next: nms-initiated
  - Notification: probe-initiated
  - UDP!

MIB-II Overview

- Relevant groups:
  - interfaces:
    - Operational state: interface ok, switched off, faulty
    - Aggregate traffic statistics: # pkts/bytes in, out,...
    - Use: obtain and manipulate operational state; sanity check (does link carry any traffic?); detect congestion
  - ip:
    - Errors: ip header error, destination address not valid, destination unknown, fragmentation problems,...
    - Forwarding tables, how was each route learned,...
    - Use: detect routing and forwarding problems, e.g., excessive fwd errors due to bogus destination addresses; obtain forwarding tables
  - egp:
    - Status information on BGP sessions
    - Use: detect interdomain routing problems, e.g., session resets
**Limitations**

- Statistics hardcoded
  - No local intelligence to: accumulate relevant information, alert NMS to prespecified conditions, etc.
- Highly aggregated traffic information
  - Aggregate link statistics
  - Cannot drill down
- Protocol: simple = dumb
  - Cannot express complex queries over MIB information in SNMPv1
    - “Fet all or nothing”
    - More expressibility in SNMPv3: expression MIB

**RMON1: Remote Monitoring**

- Advantages
  - Local intelligence & memory
  - Reduce management overhead
  - Robustness to outages
Passive Measurement Capabilities: SNMP Statistics

- Available data:
  - Summary information from and about devices

- Possible analysis:
  - (User behavior)
  - Anomaly detection (intrusion detection system)

- Disadvantages:
  - Very coarse grain information
  - (Data flood)
  - Data aggregation
  - Needle in a haystack
  - Only captures on network device information

Passive Measurement Capabilities: SNMP Statistics (2.)

- Deployment scenarios:
  - Needs cooperation of the network operator
  - Data collection/aggregation infrastructure

- Challenges
  - Hard to see all
  - Data integrity
  - Level of detail – no connection to application
  - Data correlation
  - Data privacy vs. data sharing
  - Data collection across network confederations
Passive Measurement Capabilities:
Routing Information

- Available data:
  - Summary information about devices

---

BGP Table Example (RouteViews)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Next Hop</th>
<th>AS</th>
<th>Length</th>
<th>Next Hop AS</th>
<th>Route Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>199.222.69.0</td>
<td>167.142.3.6</td>
<td>0</td>
<td>5056</td>
<td>701 7046</td>
<td>i</td>
</tr>
<tr>
<td>4.0.0.2</td>
<td></td>
<td>0</td>
<td>1</td>
<td>701 7046</td>
<td>i</td>
</tr>
<tr>
<td>204.42.253.253</td>
<td>0</td>
<td>267 2914 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>212.4.193.253</td>
<td>0</td>
<td>8918 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>205.215.45.50</td>
<td>0</td>
<td>4006 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>193.140.0.1</td>
<td>0</td>
<td>8517 9000 2548 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>165.87.32.5</td>
<td>0</td>
<td>2685 701 7046</td>
<td>e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>206.220.240.223</td>
<td>0</td>
<td>10764 1 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>203.62.248.4</td>
<td>0</td>
<td>1221 16779 1 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>203.62.252.21</td>
<td>0</td>
<td>1221 16779 1 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>157.22.9.7</td>
<td>0</td>
<td>715 1 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>193.0.0.56</td>
<td>0</td>
<td>3333 9057 3356 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>195.219.96.239</td>
<td>0</td>
<td>8297 6453 701 7046</td>
<td>i</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Prefix 199.222.69.0/24 has origin AS 7046
(whois says that 7046 is ASN-UUNET-CUSTOMER)
Passive Measurement Capabilities: Routing Information (2.)

- Available data:
  - Summary information about devices
- Possible analysis:
  - Network dynamics
  - Anomaly detection
  - Root cause analysis
- Disadvantages:
  - Very coarse grain information
  - (Data flood)
  - Data aggregation
  - Needle in a haystack
  - Only captures on network device information

Passive Measurement Capabilities: Routing Information (3.)

- Deployment scenarios:
  - Collected anyhow by the network operator (currently basis for network management)
  - Data collection/aggregation infrastructure
- Challenges
  - Lack of detail
  - Data integrity
  - Data correlation
  - (Data privacy vs. data sharing)
  - (Data collection across network confederations)
Passive Measurement: Other

- Application protocol data
- Server related data
- Access networks
- Mobile networks
- Adhoc networks
- Sensor networks
- ...

Measurements: Summary

- Challenges:
  - Scalability
    - How to reduce the amount of data to be analysed
  - Data flood
    - What to measure when the purpose is unclear
    - Expect the unexpected
  - Validation
    - How to verify any inference

- Importance:
  - Enables network management
  - Enables debugging
  - Accountability
  - Verifies presumed assumptions
Internet Dynamics: Time Scale

- **Years:** introduction of new protocols, e.g. IPv6
- **Months:** dimensioning a new circuit
- **Weeks, days:** different # of users responsible for weekly/daily cycle of traffic load, different application mix
- **Hours:** variability of traffic volume
- **Seconds:** retransmissions
- **Subseconds:** round trip times

Internet Control: Time Scale

- **Years:** IETF
- **Months:** network planning
- **Weeks:** network engineering
- **Days:** traffic engineering
- **Hours:** routing changes
- **Seconds:** TCP

Yet: User demand influences network performance but is also influenced by network performance
Terminology And General Issues

- Measurements vs. metrics
- Measurement capabilities
- Collection of measurement data
- Data reduction techniques
- Clock issues

Collection Of Measurement Data

- Need to transport measurement data
  - Produced and consumed in different systems
  - Usual scenario: large number of measurement devices, small number of aggregation points (databases)
  - Usually in-band transport of measurement data
    - Low cost & complexity
- Reliable vs. unreliable transport
  - Reliable
    - Better data quality
    - Measurement device needs to maintain state and be addressable
  - Unreliable
    - Additional measurement uncertainty due to lost measurement data
    - Measurement device can "shoot-and-forget"
Terminology And General Issues

- Measurements vs. metrics
- Measurement capabilities
- Collection of measurement data
- Data reduction techniques
- Clock issues

Controlling Measurement Overhead

- Measurement overhead
  - In some areas, could measure everything
  - Information processing not the bottleneck
  - Examples: geology, stock market, ...
  - Networking: thinning is crucial!
- Three basic methods to reduce measurement traffic
  - Filtering
  - Aggregation
  - Sampling
  - ... and combinations thereof
Filtering

Examples:
- Only record packets ...
  - matching a destination prefix (to a certain customer)
  - of a certain service class (e.g., expedited forwarding)
  - violating an ACL (access control list)
  - TCP SYN or RST packets (attacks, abandoned http download)

Aggregation

Example: identify packet flows, i.e., sequence of packets close together in time between source-destination pairs [flow measurement]
- Independent variable: source-destination
- Metric of interest: total # pkts, total # bytes, max pkt size
- Variables aggregated over: everything else

<table>
<thead>
<tr>
<th>src</th>
<th>dest</th>
<th># pkts</th>
<th># bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.b.c.d</td>
<td>m.n.o.p</td>
<td>374</td>
<td>85498</td>
</tr>
<tr>
<td>e.f.g.h</td>
<td>q.r.s.t</td>
<td>7</td>
<td>280</td>
</tr>
<tr>
<td>i.j.k.l</td>
<td>u.v.w.x</td>
<td>48</td>
<td>3465</td>
</tr>
<tr>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>
Aggregation (2.)

- Preemption: tradeoff space vs. capacity
  - Fix cache size
  - If a new aggregate (e.g., flow) arrives, preempt an existing aggregate
    - For example, least recently used (LRU)
  - Advantage: smaller cache
  - Disadvantage: more measurement traffic
  - Works well for processes with temporal locality
    - Because often, LRU aggregate will not be accessed in the future anyway → no penalty in preemption

Sampling

- Examples
  - Systematic sampling
    - Pick out every 100th packet and record entire packet/record header
    - Ok only if no periodic component in process
  - Random sampling
    - Flip a coin for every packet, sample with prob. 1/100
  - Record a link load every n seconds
Sampling (2.)

- What can we infer from samples?
  - Easy:
    - Metrics directly over variables of interest, e.g., mean, variance etc.
    - Confidence interval = “error bar”
      - Decreases as $1/\sqrt{n}$
  - Hard:
    - Small probabilities: “Number of SYN packets sent from A to B”
    - Events such as: “has X received any packets”?

Sampling (3.)

- Hard:
  - Metrics over sequences
  - Example: “how often is a packet from X followed immediately by another packet from X?”
    - Higher-order events: probability of sampling $i$ successive records is $p^i$
    - Would have to sample different events, e.g., flip coin, then record $k$ packets
Sampling (4.)

- Sampling objects with different weights

- Example:
  - Weight = flow size
  - Estimate average flow size
  - Problem: a small number of large flows can contribute very significantly to the estimator

- Stratified sampling: make sampling probability depend on weight
  - Sample "per byte" rather than "per flow"
  - Try not to miss the "heavy hitters" (heavy-tailed size distribution!)

\[ p(x) \text{ constant} \]
\[ p(x) \text{ increasing} \]

Sampling (5.)

Object size distribution

Estimated mean:

\[ \hat{\mu} = \frac{1}{n} \sum x \cdot n(x) \]

\[ r(x) = \text{# samples of size } x \]

\[ x \cdot r(x) : \text{contribution to mean estimator} \]

Variance mainly due to large \( x \)

Better estimator: reduce variance by increasing # samples of large objects
Basic Properties

<table>
<thead>
<tr>
<th></th>
<th>Filtering</th>
<th>Aggregation</th>
<th>Sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>exact</td>
<td>exact</td>
<td>approximate</td>
</tr>
<tr>
<td>Generality</td>
<td>constrained a-priori</td>
<td>constrained a-priori</td>
<td>general</td>
</tr>
<tr>
<td>Local Processing</td>
<td>filter criterion for every object</td>
<td>table update for every object</td>
<td>only sampling decision</td>
</tr>
<tr>
<td>Local memory</td>
<td>none</td>
<td>one bin per value of interest</td>
<td>none</td>
</tr>
<tr>
<td>Compression</td>
<td>depends on data</td>
<td>depends on data</td>
<td>controlled</td>
</tr>
</tbody>
</table>

Combinations

- In practice, rich set of combinations of filtering, aggregation, sampling
- Examples:
  - Filter traffic of a particular type, sample packets
  - Sample packets, then filter
  - Aggregate packets between different source-destination pairs, sample resulting records
  - When sampling a packet, sample also \( k \) packets immediately following it, aggregate some metric over these \( k \) packets
  - ... etc.
Terminology And General Issues

- Measurements vs. metrics
- Measurement capabilities
- Collection of measurement data
- Data reduction techniques
- Clock issues

Clock Issues

- Time measurements
  - Packet delays: we do not have a “chronograph” that can travel with the packet
    - Delays always measured as clock differences
  - Timestamps: matching up different measurements
    - E.g., correlating alarms originating at different network elements

- Clock model:
  \[ T(t) = T(t_0) + R(t_0)(t - t_0) + \frac{1}{2} D(t_0)(t - t_0)^2 + O((t - t_0)^3) \]
  
  \( T(t) \): clock value at time \( t \)
  \( R(t) \): clock skew : first derivative
  \( D(t) \): clock drift : second derivative
**Delay Measurements: Single Clock**

- Example: round-trip time (RTT)
- $T_1(t_1) - T_1(t_0)$
- only need clock to run approx. at the right speed

**Delay Measurements: Two Clocks**

- Example: one-way delay
- $T_2(t_1) - T_1(t_0)$
- very sensitive to clock skew and drift
Clock Issues (2.)

- **Time-bases**
  - **NTP (Network Time Protocol):** distributed synchronization
    - No additional hardware needed
    - Not very precise and sensitive to network conditions
    - Clock adjustment in “jumps” → switch off before experiment!
  - **GPS**
    - Very precise (100ns)
    - Requires outside antenna with visibility of several satellites
  - **SONET clocks**
    - In principle available & very precise

NTP: Network Time Protocol

- **Goal:** disseminate time information through network
- **Problems:**
  - Network delay and delay jitter
  - Constrained outdegree of master clocks
- **Solutions:**
  - Use diverse network paths
  - Disseminate in a hierarchy (stratum $i$ → stratum $i+1$)
  - A stratum-$i$ peer combines measurements from stratum $i$ and other stratum $i-1$ peers
NTP: Peer Measurement

Message exchange between peers
- clock 2 knows \([T_2(t_1), T_1(t_2), T_1(t_3)]\) at \(t_4\)
- assuming \(t_2 - t_1 \approx t_4 - t_3\),
  
  offset \(\approx \frac{T_1(t_2) + T_1(t_3) - T_2(t_1) - T_2(t_3)}{2}\)
  
  roundtrip delay \(\approx T_1(t_2) - T_1(t_3) - T_2(t_1) + T_2(t_3)\)

NTP: Combining Measurements

Clock filter
- Temporally smooth estimates from a given peer

Clock selection
- Select subset of “mutually agreeing” clocks
  - Intersection algorithm: eliminate outliers
  - Clustering: pick good estimates (low stratum, low jitter)

Clock combining
- Combine into a single estimate
NTP: Status And Limitations

- Widespread deployment
  - Supported in most OSs, routers
  - >100k peers
  - Public stratum 1 and 2 servers carefully controlled, fed by atomic clocks, GPS receivers, etc.

- Precision inherently limited by network
  - Random queueing delay, OS issues...
  - Asymmetric paths
  - Achievable precision: O(20 ms)