User-level Internet Path Diagnosis

Ratul Mahajan
Neil Spring
David Wetherall
Thomas Anderson

University of Washington
Diagnosing performance of Internet paths is hard

The Internet as a black box

- Multiple administrative domains
  - operators may be equally clueless
- Policy routing
  - asymmetric paths (round trip tools such as ping don’t work well)
  - path to intermediate routers may not be a prefix of the end-to-end path to the destination
- Performance may depend on the application
  - packet size, inter-packet spacing, protocol, port number, …..
Our goal is “user-level” diagnosis

- **Diagnosis:** identify and localize performance faults that impact applications
  - loss, reordering, queuing delay, …

- **User-level:** without privileged access to routers
  - useful for both end users and network operators

- Diagnosis is useful (even if you cannot fix yourself)
  - transparency will lead to faster problem resolution
  - intelligently route around the fault
Existing diagnosis tools have limitations

- `ping/traceroute/pathchar` measure round trip path to routers
  - path asymmetry conflates forward and reverse paths

- Effective diagnosis requires router support beyond packet reflection
Approach and outline

◆ Architecture
  • what minimal support is needed to enable user-level diagnosis in Internet-like networks?

◆ Build practical tools
  • tulip

◆ Explore Internet evolution to improve diagnostic support
An architecture for path diagnosis

Start with an ideal solution
- routers log all packets they forward
- users diagnose their paths through trace analysis
- complete but impractical

Reduce to a practical architecture
1. all routers on the path embed diagnostic info in packets
   - timing, flow counters, and path information
2. the source samples one router to embed diagnostic info
An architecture for path diagnosis (2)

- Lightweight, in-band packet marking
  - almost as powerful as the complete path trace

<table>
<thead>
<tr>
<th>field</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>sampler</td>
<td>selects the sampling router</td>
</tr>
<tr>
<td>timestamp</td>
<td>local time at the sampling router</td>
</tr>
<tr>
<td>flow counter</td>
<td># of pkts processed for this flow</td>
</tr>
<tr>
<td>path signature</td>
<td>to detect path changes</td>
</tr>
</tbody>
</table>

- Timing, flow counters and path information provide effective diagnostic support
Approach and outline

◆ Architecture
  • what minimal support is needed to enable user-level diagnosis in Internet-like networks?

◆ Build practical tools
  • tulip

◆ Explore Internet evolution to improve diagnostic support
Overview of tulip

Measuring forward path to routers is the basic building block

- Localizes reordering, queuing and loss (so far)
  - single-ended: works from a host to an arbitrary IP address

- Infers link properties by subtracting path properties
  - path to router should be a prefix of the end-to-end forward path
ICMP timestamps are used to access router’s clock [cing]
- 1 ms resolution; supported by over 90% routers
- prefix path property may not hold

Queuing inferred from delay variation

Engineering – clock calibration, response generation time
Loss on the forward path

- Loss measurements use the IP identifier field in IP packets
  - over 70% of routers implement IP-ID as a counter
  - common counter for all probing sources
- Unambiguous detection of forward path loss for data packets
  - when control responses get consecutive IP-IDs
- Robust to response rate-limiting at the routers
Experimental evaluation of tulip

◆ What is the resolution of fault localization?
  • diagnosis granularity

◆ Is it accurate?
  • end-to-end correctness
  • consistency (monotonic increase along the path)
Granularity: uncertainty in the location of the fault

- when a router does not support the required features
- when probes take a non-prefix path to a router
Diagnosis granularity of tulip (2)

- Median is 2 hops for loss and 4 hops for queuing
  - ICMP timestamp probes do not have the prefix path property

- Round trip probing can further improve diagnosis granularity
Experimental evaluation of tulip

◆ What is the resolution of fault localization?
  • diagnosis granularity

◆ Is it accurate?
  • end-to-end correctness
  • internal consistency (monotonic increase along the path)
Consistency along the path (queuing)

Tulip’s one-way measurements are consistent
Round trip measurements are polluted by reverse path conflation
Consistency along the path (queuing)

- Tulip’s one-way measurements are consistent
- Round trip measurements are polluted by reverse path conflation

queuing delta = delay at the far end – delay at the near end
Tulip in action

Tulip can help build more scalable network monitoring and overlay routing systems

ratul | sosp | 2003
Approach and outline

- **Architecture**
  - what minimal support is needed to enable user-level diagnosis in Internet-like networks?

- **Build practical tools**
  - tulip – a tool to diagnose reordering, loss, and queuing delay

- **Explore Internet evolution to improve diagnostic support**
Recall: an architecture for path diagnosis

- Lightweight, in-band packet marking
  - almost as powerful as the complete path trace

<table>
<thead>
<tr>
<th>field</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>sampler</td>
<td>selects the sampling router</td>
</tr>
<tr>
<td>timestamp</td>
<td>local time at the sampling router</td>
</tr>
<tr>
<td>flow counter</td>
<td># of pkts processed for this flow</td>
</tr>
<tr>
<td>path signature</td>
<td>to detect path changes</td>
</tr>
</tbody>
</table>

- Timing, flow counters and path information provide effective diagnostic support
Tulip approximates the architecture in the Internet

<table>
<thead>
<tr>
<th>sampler</th>
<th>TTL or router’s IP addr</th>
</tr>
</thead>
<tbody>
<tr>
<td>timestamp</td>
<td>ICMP timestamps</td>
</tr>
<tr>
<td>flow counter</td>
<td>IP-ID counter</td>
</tr>
<tr>
<td>path signature</td>
<td>hop count</td>
</tr>
</tbody>
</table>

- Approximations (and tulip) have limitations
  - measurement probes are out-of-band
  - ICMP timestamp issues (next slide)
  - IP-ID counter is shared
  - path changes can go undetected
- Moving the Internet towards the architecture improves diagnostic support
  - identify small changes with big benefits
Better timing information

- **Problems:**
  - timing information is separate from flow counters
  - ICMP timestamps require directly addressing the router
    - routing issues reduces their value

- **Simple fix: timestamp TTL-expired messages**
  - backwards compatible, incrementally deployable
    - use 32 unused bits in the TTL-expired messages
Better counter support

- **Problem:**
  - IP-ID is a shared counter
    - what if all of you start using tulip?
    - the architecture suggests per-flow counters

- **Simple fix:** maintain N (constant) counters
  - hash source address and probe IP-ID to pick the counter
  - backwards compatible, incrementally deployable (today, N=1)
Summary

◆ Tulip enables end users to diagnose Internet paths
  • co-opts router support by exploiting well-deployed router features

◆ Architectural arguments:
  • features used by tulip approximate a lightweight architecture for user-level path diagnosis
  • approximations suggest evolutionary changes to improve Internet’s diagnostic support

◆ Future work: extend tulip with
  • tomography to improve diagnosis granularity
  • higher layer protocol diagnosis