Effective Diagnostic Strategies for Wide Area Networks

Matt Mathis <mathis@psc.edu>
Peter O’Neil <poneil@ucar.edu>

11-Dec-2003

Slides: http://www.psc.edu/~mathis/papers/Diagnosis200312/
New Diagnostic Research

- Freshly funded NSF STI
  - 3 Years
  - PSC + NCAR
  - Grant ANI-0334061

- Builds on Web100 & Net100
  - Including some existing prototypes

- Addresses the next tier of performance problems
  - Web100 was "End-system only"
  - Now turn our attention to the path
Key Observation

- Symptoms are scaled by RTT
  - For server S
  - Local Client, LC, works fine
  - Remote Client, RC, fails

- Applies to:
  - most types of flaws
Examples

- Chatty application (e.g. 50 RTT per user request)
  - On a 1 ms LAN, 50 ms total time
  - On a 100 ms WAN, 5 s total time

- Fixed TCP buffer space (e.g. 32 kBytes)
  - On a 1 ms LAN, 200 Mb/s limit
  - On a 100 ms WAN, 2 Mb/s limit

- Packet Loss (e.g. 1%, with 9 kB packets)
  - On a 1 ms LAN, 500 Mb/s limit
  - On a 100 ms WAN, 5 Mb/s limit

- Also applies to TCP stack bugs
  - Defective Fast Retransmit, RTO timers, No SACK, etc
Confounds classic diagnostics

- False pass on short paths
  - Obscures true local bugs
  - Incorrectly implicates other components

- Stymies end-to-end diagnosis
  - e.g. In situ tomography
Our primary approach

- Web100 based diagnostic server
  - Simple TCP test to a test target
  - Use MIB & model to rescale results

- TCP discard server for test target
  - Trivial to widely deploy
  - C or Java

- Estimates results for long path
End to End Diagnosis

- Approx 1 Diagnostic Server per campus/Gigapop/Backbone/Interconnect

- Test Targets at almost every pop/hub or LAN (or above)
End user tool

- Web/java client
  - With built in test target
  - Invokes test on DS back to self


- http://whisper.cs.utk.edu:7123/
Other approaches in the project

- Bench test applications and end-systems (stacks)
  - NOTE: out of scope for this meeting!

- Use long ideal (virtual) paths
  - Dummynet style emulated delay
  - Tunnel or VPN style "scenic" routing

- ..... etc
Goal

- Tools to compensate for results that scale by RTT
Our non-goals

- Overall diagnostic framework or architecture
- New authentication and authorization
- Locating diagnostic servers and test targets
- Global infrastructure
Epilogue

- All single point failures are easy to find and correct

- Remaining failures are co-failures:
  - RTT, loss and MTU
  - RTT and application design
  - Packet rate limit and MTU
  - Queue size and burstyness due to
    - ACK compression and cross traffic or
    - Application design

- Failures are due to interaction between elements
  - THERE IS NO INTRINSIC BLAME