A Reactive Measurement Framework

Mark Allman, Vern Paxson
International Computer Science Institute

Passive and Active Measurement Conference
April 2008

*"This town, buddy, has done it's share of shovin'
This town taught me that it's never too late"*

Overview

- Difficult to understand networks due to the vast array of integrated components
  - some known, some not (whee!)
- Usual approach is to *measure and wonder*

Overview (cont.)

- E.g., we know nothing about the routing state by looking at a packet trace
- E.g., we know nothing about the connectivity when observing a DNS lookup failure
- E.g., we don't understand if a web fetch failed because of DDoS, duplex mis-match, proxy failure, ...
- E.g., we don't understand why a SYN didn't elicit a SYN+ACK

Reactive Measurement

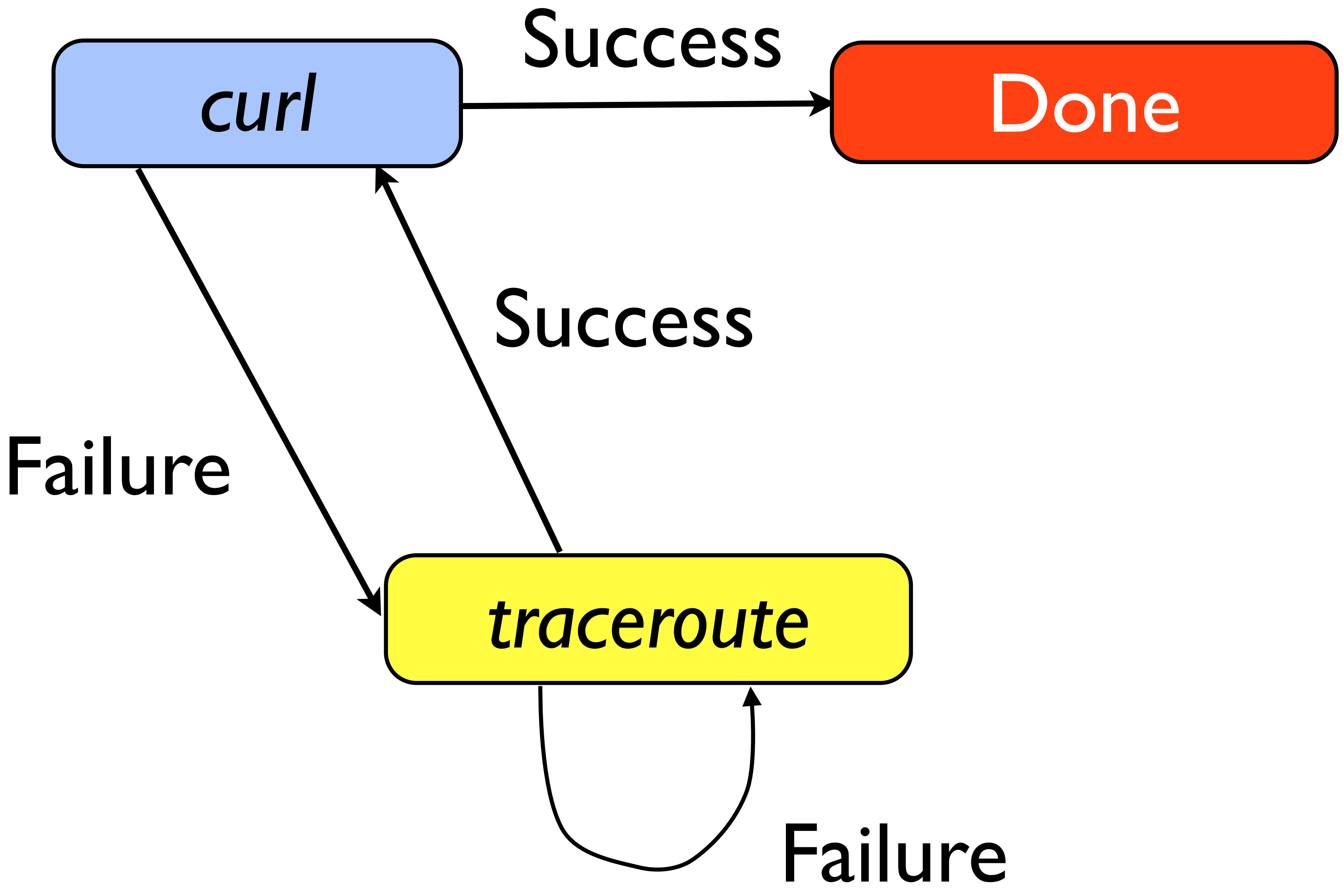
- Old way: measurement as an *event*
- New way: measurement as a *process*

REM (cont.)

- REM calls for disparate measurement tools to be orchestrated in a way that leads to better fundamental understanding

- Simple but powerful notion

REM (cont.)



Related Work

- Used operationally
  - E.g., SNMP traps
  - E.g., IDS
- Ad-hoc use in research

Application #1

- Fundamental new approach to answering questions
    - E.g., how long do DNS failures persist?

Application #2

- Targeted measurements
  - E.g., packet capture only at key times
  - Eases measurement logistics
    - *not a small contribution*

Application #3

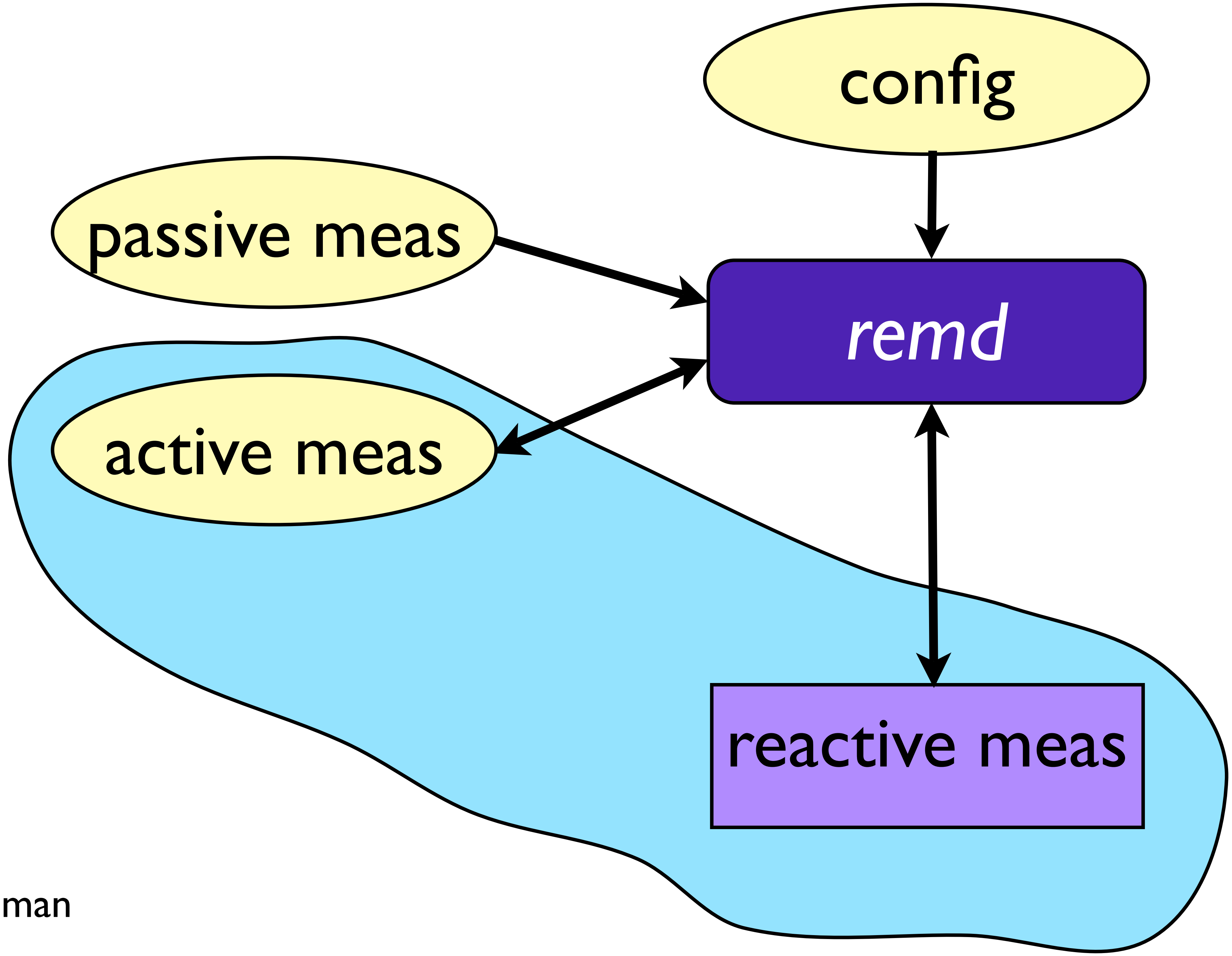
- Network anomalies can be better understood
    - now often a footnote in a paper

A REM Architecture

- Goal: to build a simple *glue* to allow for ...
  - quickly bonding disparate measurements
  - dealing with general resource issues (e.g., runaway measurements)
  - dealing with collecting data intelligently

- Hope: a reasonable toolkit will *foster* REM studies

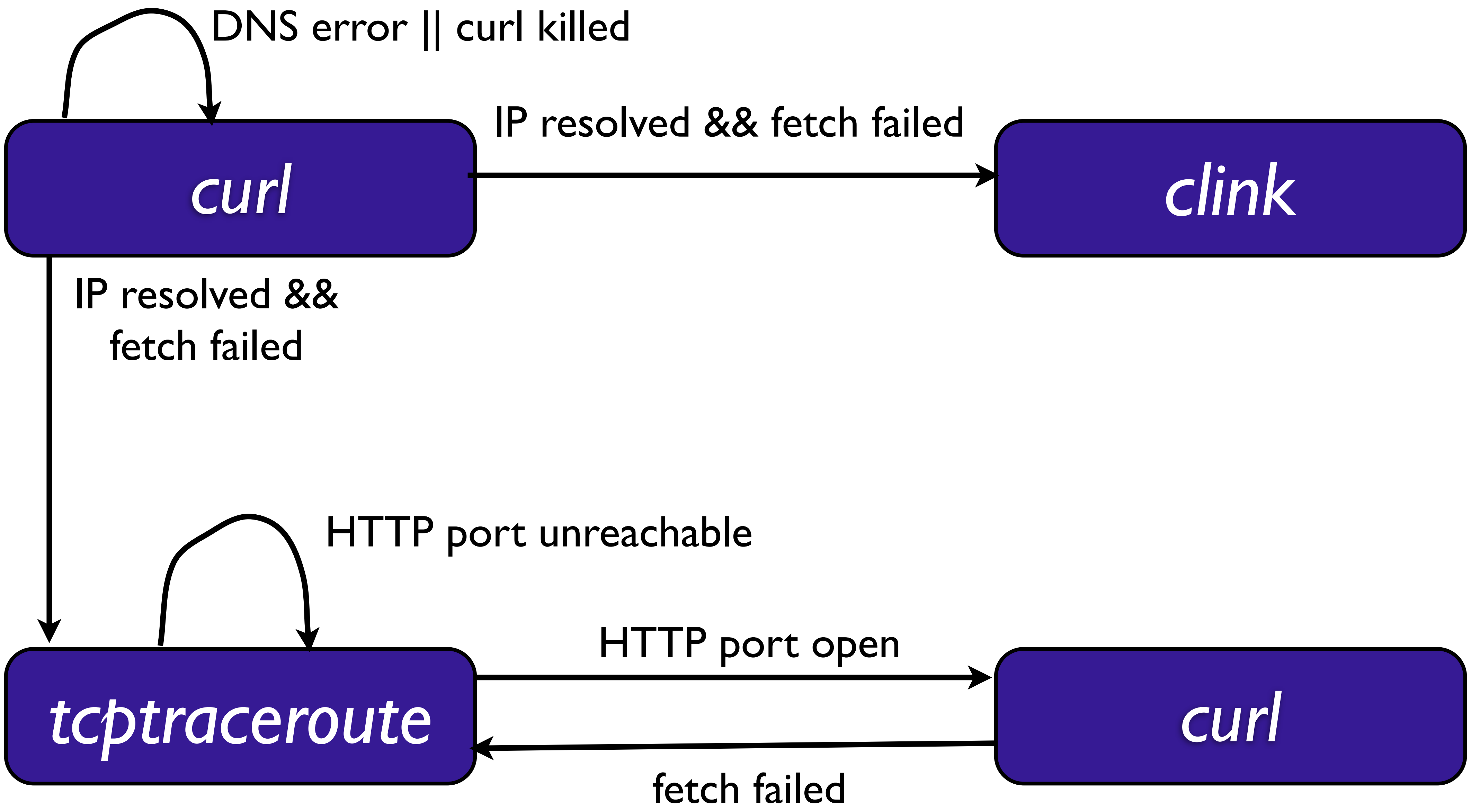
REM Architecture (cont.)



Sample Experiment

- What is the cause of web fetch failures?

Experiment (cont.)



Experiment (cont.)

- 408K fetches, Dec/07--Feb/08
- 5.5K failures on initial curl (1.4%)
  - 23% were fixed within the measurement period (~20 min)
  - 77% persisted throughout the measurement period

Experiment (cont.)

- Initial failures that ultimately succeed
  - 49%: connected, but failed to get data
  - 18%: DNS error
  - 15%: successful HTTP transaction, but no actual content
  - 10%: failure to connect to server
  - 5%: partial fetch completed
  - etc.

Experiment (cont.)

- Persistent failures
  - 50%: DNS errors
  - 30%: connected, but failed to get data
  - 11%: failure to connect to server
  - 6%: successful HTTP transaction, but no actual content
  - etc.

Problem: State Machines

- We have found state machines to be somewhat restrictive
  - E.g., simple cases where we want to assess the network in two ways and then take action based on both results
  - E.g., engaing with external measurement infrastructure such as DipZoom
- Lousy workarounds:
  - Run serially
  - Push complexity to wrapper scripts

Contributions

- Looking to examine and espouse the general power of the REM approach

- Building a toolkit to make utilizing REM techniques straightforward

- Feedback on our initial thinking.

INTERNATIONAL
COMPUTER SCIENCE
INSTITUTE

“If there’s something you want, If there’s something you need, ...”

Mark Allman

mallman@icir.org
http://www.icir.org/mallman/