

Request Tracking in Server Systems (cont.)

Kai Shen

Dept. of Computer Science, University of Rochester

Magpie [Barham et al. OSDI2004]

- Tracking requests transparently
 - dumping bunch of system events
 - analyze the events afterward to find out request context propagations
 - process A sends something into a socket and process B reads out of the same socket a bit later
 - an interrupt handler is linked to a page cache entry which is later read by process C
- Flexibility
 - a framework with user-defined rules
- Extending this into application level
 - user-level synchronization with explicit tracing
- Not on-the-fly
 - why not?

2/5/2008

URCS 573 - Spring 2008

2

Utilization of Magpie Request Tracking

- Request classification
 - each request is identified by its events
 - tree or chain?
 - put similar ones into groups/clusters
 - what is the good metric to quantify the difference between requests?
- Request modeling
 - given an online workload, you can say
 - 80% are like this ...
 - 19% are like this ...
 - a couple of really weird ones

2/5/2008

URCS 573 - Spring 2008

3

On-the-Fly Transparent Request Tracking [URCS work, ASPLOS2008]

- On-the-fly tracking allows immediate system reaction
 - adaptive OS management policy
 - anomaly detection and quarantine
- Most request propagations are through inter-processing communications
 - tag request contexts to messages
 - problem with tagging message
 - message may be asynchronous
 - message boundaries in stream-based communications
 - more problems?
- Other request propagations?

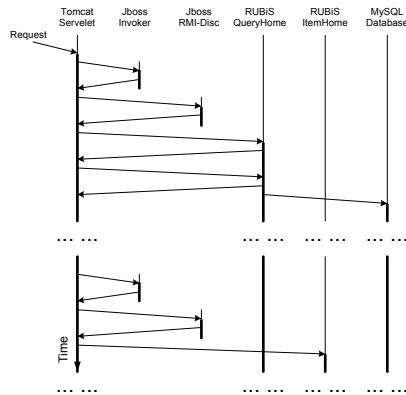
2/5/2008

URCS 573 - Spring 2008

4

An Example of Request Propagation

- Server request context propagation [with help from Chris Stewart]



2/5/2008

URCS 573 - Spring 2008

5

Distributed Request Tracking

- Identify request contexts across machines
- Overhead of passing tags and distributed analysis
- Issues related to distributed management

2/5/2008

URCS 573 - Spring 2008

6

Discussions

- Request tracking mechanisms
 - transparency
 - on-the-fly vs. offline
 - flexibility
 - robustness
 - support for distributed tracking
- Utilization of request tracking
 - request-granularity resource allocation
 - workload modeling
 - request-granularity OS management, anomaly detection

2/5/2008

URCS 573 - Spring 2008

7