

# Request Tracking in Server Systems

Kai Shen

Dept. of Computer Science, University of Rochester

## Macroscopic/Microscopic Views

- Stochastic queueing analysis/operational laws
  - macroscopic view
- Tracking individual request
  - microscopic view
  - also allows management at request granularity
  - a lot more intrusive now - ideally only changing operating system
- What is a "request"?
  - from application point of view, a task the server fulfills on behalf of a user demand;
  - from system point of view, the set of all activities (in the server system) in fulfilling a single application request

1/31/2008

URCS 573 - Spring 2008

2

## Track Request CPU Usage

- Oftentimes a request maps to a process/thread
- How to track a process CPU usage in OS?
  - sampling vs. tracking context switches
  - accounting for interrupt handlers
  - accounting for background tasks
- Exceptions that a request does not exactly map to a process
  - process pooling
  - multi-stage server

1/31/2008

URCS 573 - Spring 2008

3

## Resource Containers [Banga et al. OSDI1999]

- Tracking requests
  - applications pass request handles when requests go from one stage to the next
- Request-granularity resource management
  - what is the default scheduling unit in operating systems?
  - decouple scheduling context, protection domain, and resource allocation context
  - flexible/better resource control - e.g., give certain users higher priority or guaranteed CPU allocation

1/31/2008

URCS 573 - Spring 2008

4