Value-added Services on Software-Programmable Routers

Systems Software and Architecture Laboratory

Department of Computer Sciences

http://www.cs.purdue.edu/people/yau

http://ssal.cs.purdue.edu/

Motivations

- More sophisticated network contents
- More demanding network users
- Value-added services
  - accounting
  - security (copyright, authentication)
  - active caching
  - ...

Challenges

- Heterogeneous users
  - needs, priorities, purchased shares
- Untrusted programs
  - greedy, buggy, malicious, ...
- Diverse resources
  - space-shared, time-shared
- Diverse resource *bindings*

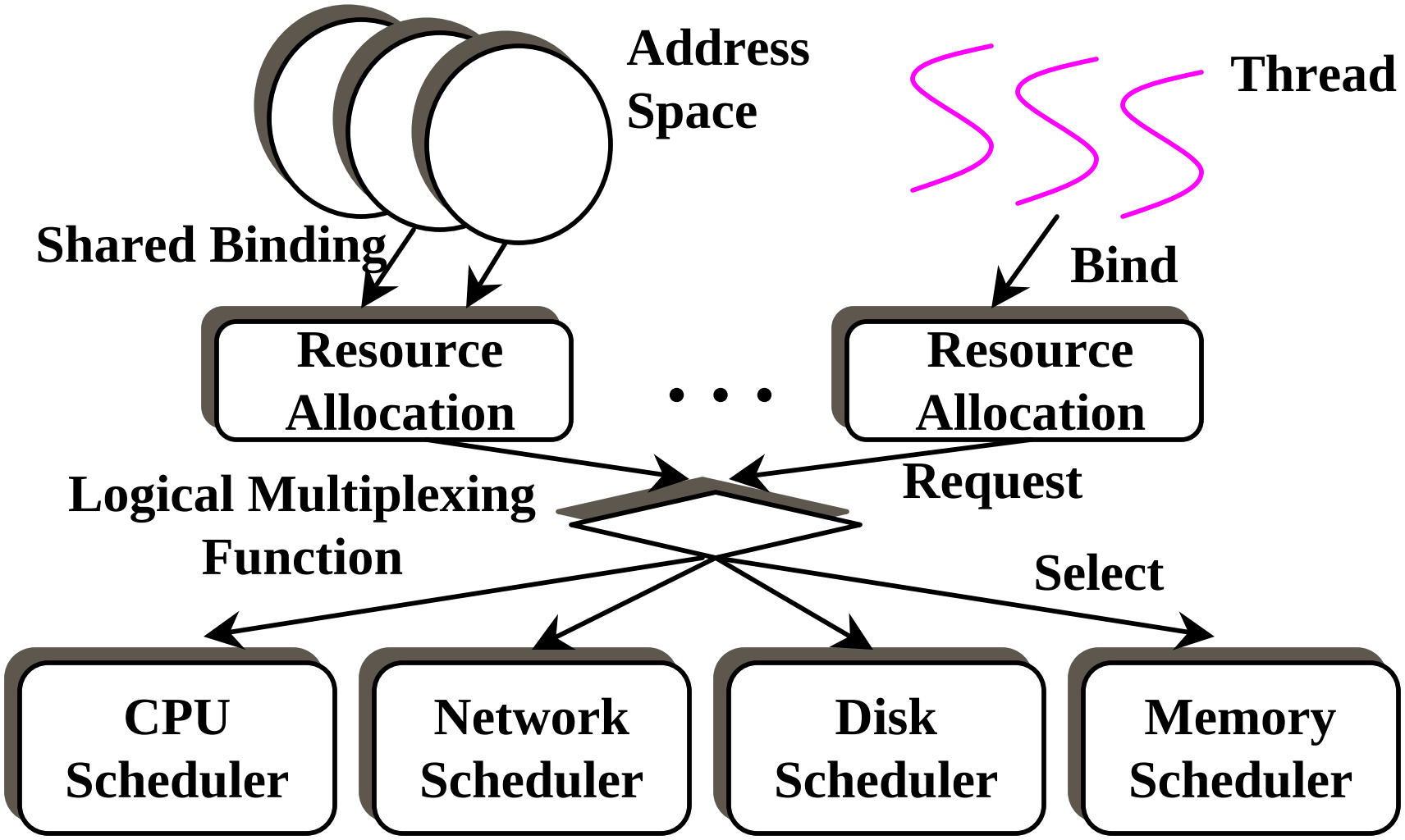
Our Approach

- **Virtualized** router resources
  - virtual machines
- **Orthogonal** fine-grained allocations
  - Resource Allocation objects
- Flexible/scalable **packet classification**
  - resource binding, per-flow processing
- Efficiency, modularity, configurability

Resource Abstraction

- Kernel *Resource Allocation* objects
- Independent/orthogonal objects
  - relative to resource consumers
- Flexible bindings to resource consumers
  - shared binding, dynamic binding (with run-time information), configurable parameters
- Hierarchical Scheduling of multiple resource types
  - CPU, network, memory pool, disk bandwidth

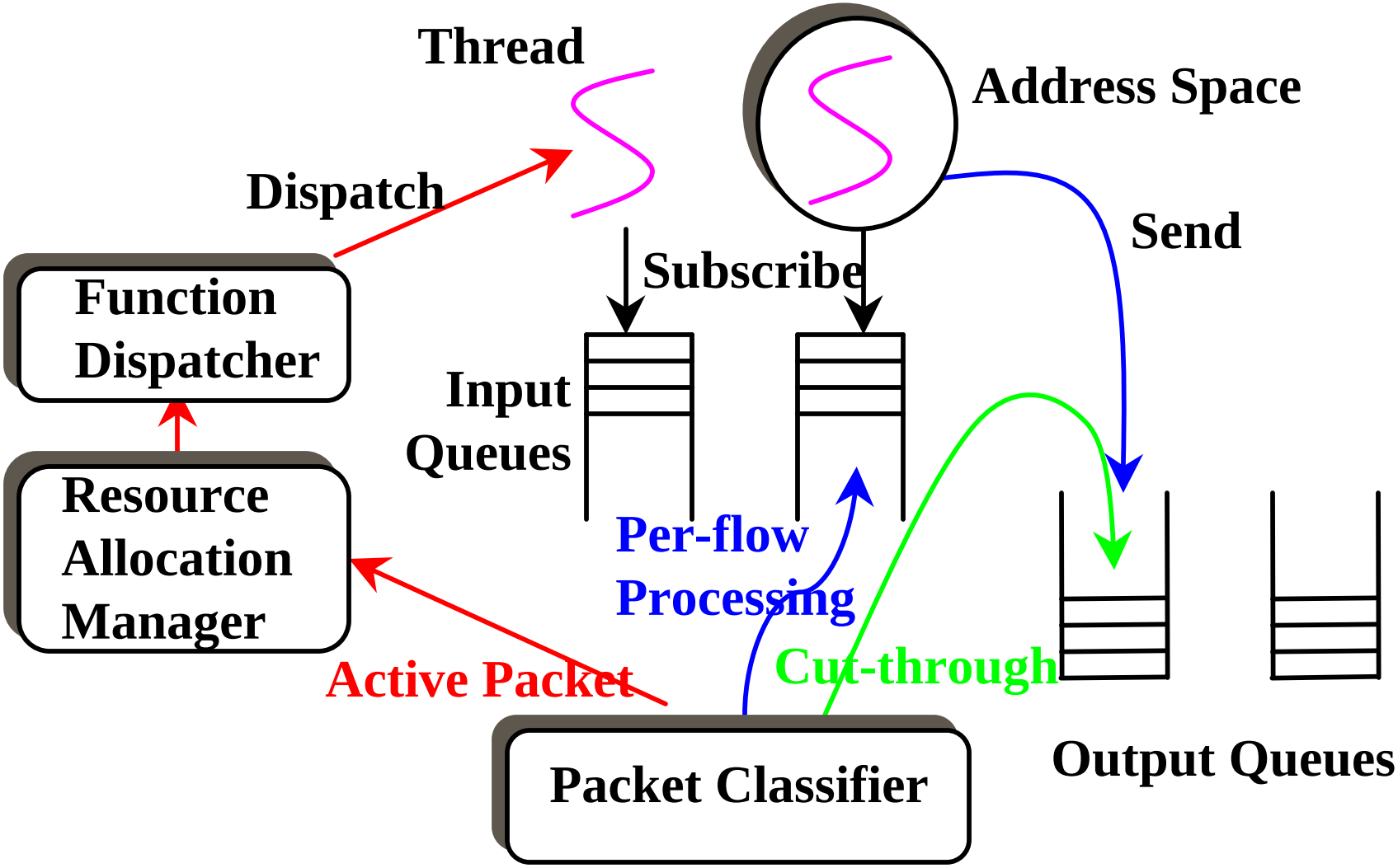
Schedulers for Resource Allocations



Packet Forwarding

- Possibilities
  - active program dispatch
    - trusted (kernel thread), untrusted (user process)
  - Per-flow processing
    - subscribed by dispatched router programs
    - security processing, application-level routing
  - Cut-through fast path
    - minimal delay

Processes in the router



Packet forwarding decision

- Based on packet header information
- Packet classification
  - scalable to many dimensions
  - scalable to many classification rules
  - flexible
    - support multiple and least-cost matches

Resource Binding Decision

- Active packet starts router program
- Program must run with resource allocation
  - *Which* allocation?
  - Retrieved as part of packet classification
  - Request to create new allocation
  - Request to use existing allocation with given key

System Implementation

- Extension to Solaris 2.5.1
- Deployed on UltraSPARC/Pentium network
  - Ethernet, Fast Ethernet, Myrinet
  - Support for **existing** applications
- Modular subsystems with well-defined interfaces
- Simple command interfaces to launch legacy applications

Basic Costs

- Resource Allocation control
  - create + delete 15.4 microseconds kernel, 19.6 user
  - bind 4.8 kernel, 9.0 user
  - unbind 2.4 kernel, 6.6 user
- Function dispatch
  - thread: about 145 microseconds, low variance
  - process: 0.77 to 1.1 ms, application-dependent

Packet Forwarding Performance

- Five dimension
  - exact, prefix, range, wildcard
- Database size up to 256 K rules
- Average lookup cost of 7.8 microseconds
  - 1.1 Gb/s throughput for 1000 byte packets
- Add/delete 10.8/14.9 microseconds
  - 67,000 updates per second

Summary

- Resource management important for software-programmable routers
- Building system prototype as solution step
  - packet classification
  - router program dispatch
  - unified and orthogonal resource abstraction
  - schedulers for major resource types