CS 268: Lecture 11 (Differentiated Services)

Ion Stoica

March 6, 2001

Administrative Announcement

- Next Monday (March 11) project presentations:
- Each group has 8 minutes
  - 5 minutes for presentations
  - 3 minutes for questions
- Time will be very strictly enforced
- Don’t use more than five slides (including the title slide)

Presentation

- 1st slide: Title
- 2nd slide: motivations and problem formulation
  - Why is the problem important?
  - What is challenging/hard about your problem
- 3rd slide: main idea of your solution
- 4th slide: status
- 5th slide: future plans and schedule

What is the Problem?

- Goal: provide support for wide variety of applications:
  - Interactive TV, IP telephony, on-line gamming (distributed simulations), VPNs, etc
- Problem:
  - Best-effort cannot do it (see previous lecture)
  - Intserv can support all these applications, but
    - Too complex
    - Not scalable

Differentiated Services (Diffserv)

- Build around the concept of domain
- Domain – a contiguous region of network under the same administrative ownership
- Differentiate between edge and core routers
- Edge routers
  - Perform per aggregate shaping or policing
  - Mark packets with a small number of bits; each bit encoding represents a class (subclass)
- Core routers
  - Process packets based on packet marking
- Far more scalable than Intserv, but provides weaker services

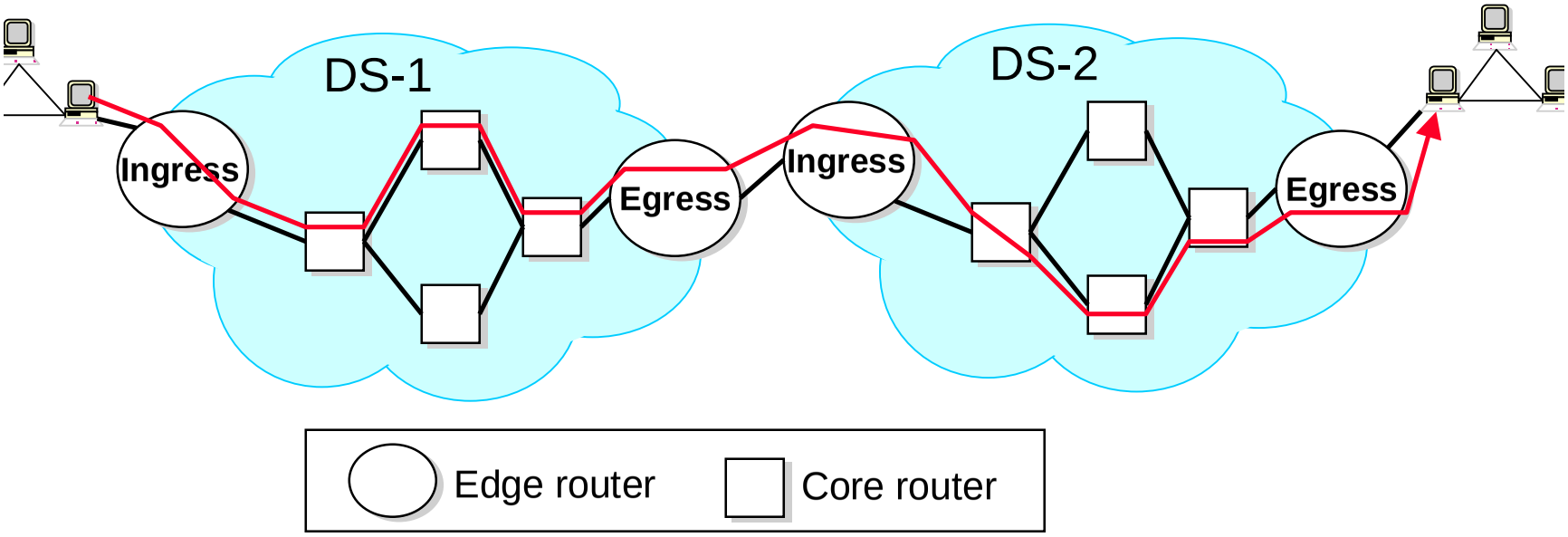
Diffserv Architecture

Ingress routers

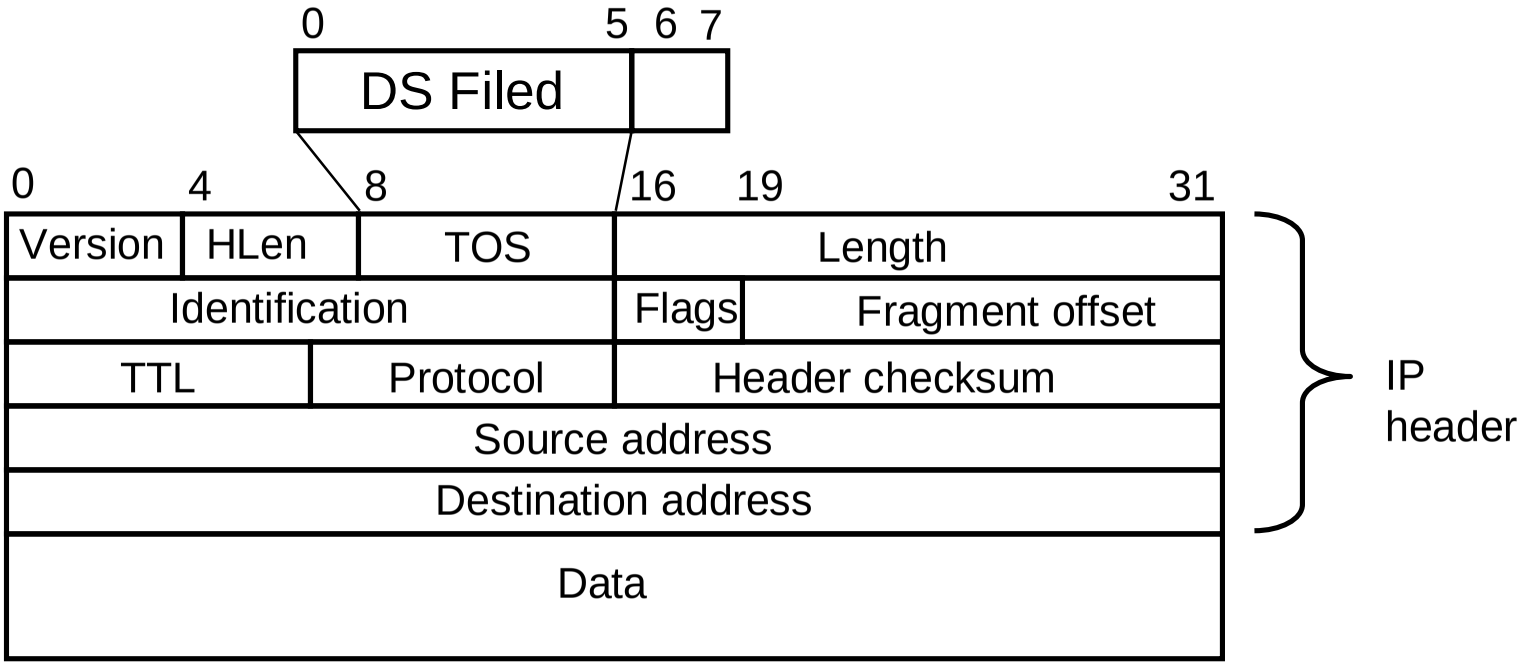
- Police/shape traffic
- Set Differentiated Service Code Point (DSCP) in Diffserv (DS) field

Core routers

- Implement Per Hop Behavior (PHB) for each DSCP
- Process packets based on DSCP



Differentiated Service (DS) Field



- DS filed reuse the first 6 bits from the former Type of Service (TOS) byte
- The other two bits are proposed to be used by ECN

Differentiated Services

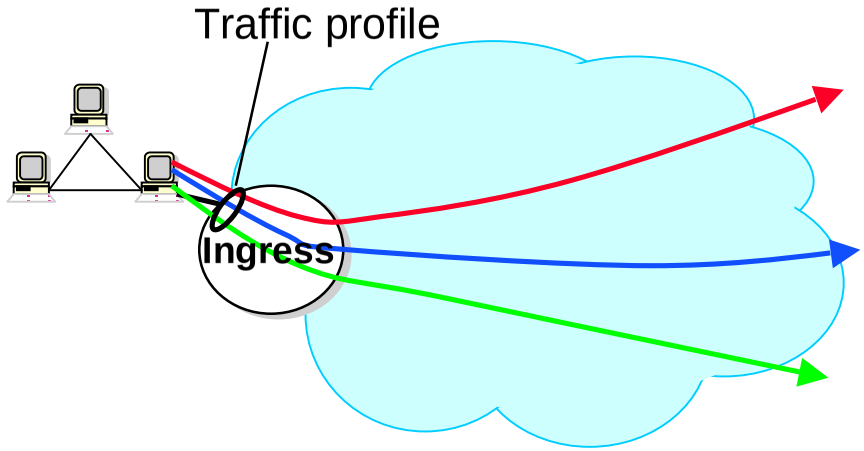
- Two types of service
  - Assured service
  - Premium service
- Plus, best-effort service

Assured Service [Clark & Wroclawski '97]

- Defined in terms of user profile, how much assured traffic is a user allowed to inject into the network
- Network: provides a lower loss rate than best-effort
  - In case of congestion best-effort packets are dropped first
- User: sends no more assured traffic than its profile
  - If it sends more, the excess traffic is converted to best-effort

Assured Service

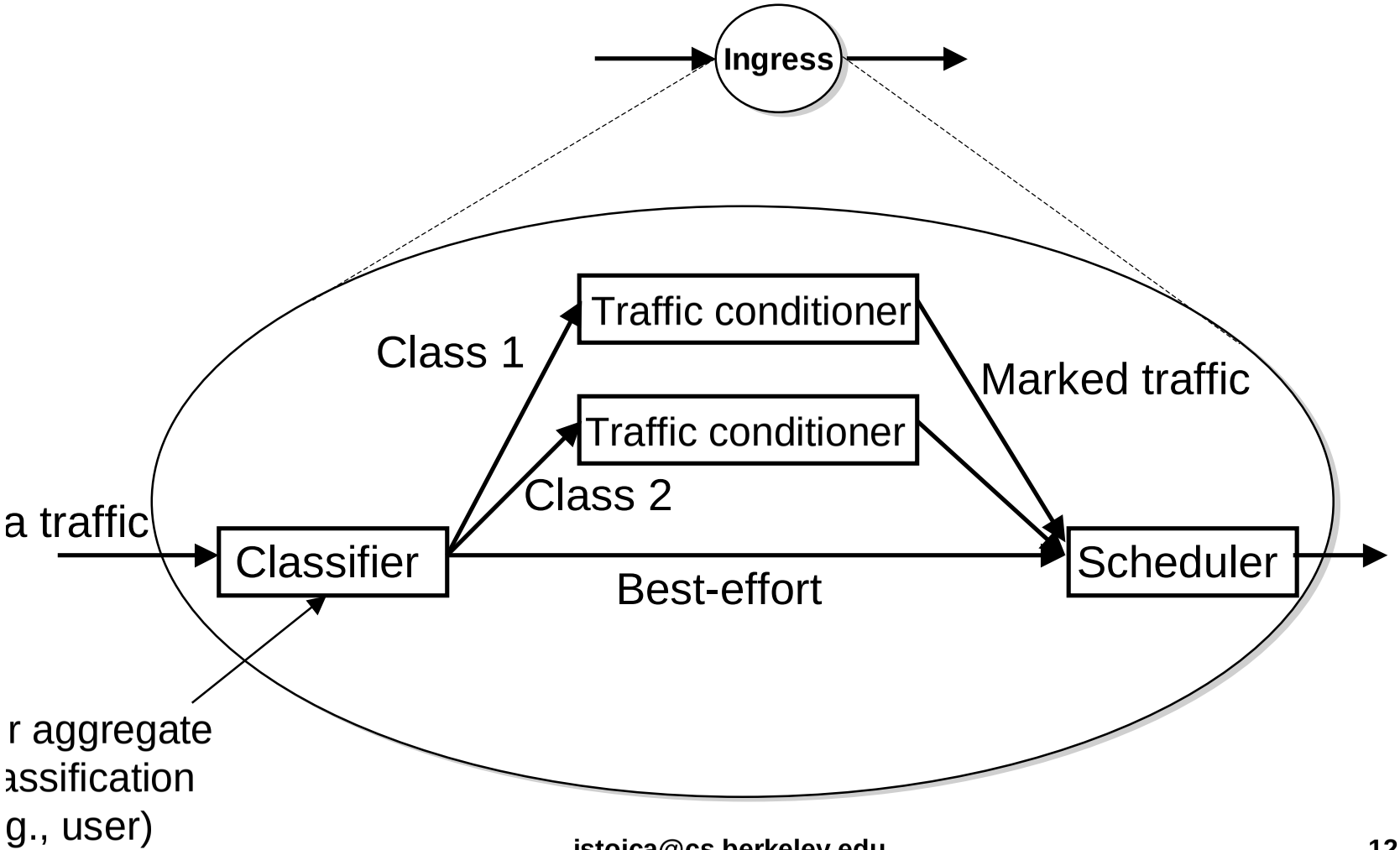
- Large spatial granularity service
- Theoretically, user profile is defined irrespective of destination
  - All other services we learnt are end-to-end, i.e., we know destination(s) apriori
- This makes service very useful, but hard to provision (why ?)



Premium Service [Jacobson '97]

- Provides the abstraction of a virtual pipe between an ingress and an egress router
- Network: guarantees that premium packets are not dropped and they experience low delay
- User: does not send more than the size of the pipe
  - If it sends more, excess traffic is delayed, and dropped when buffer overflows

Edge Router

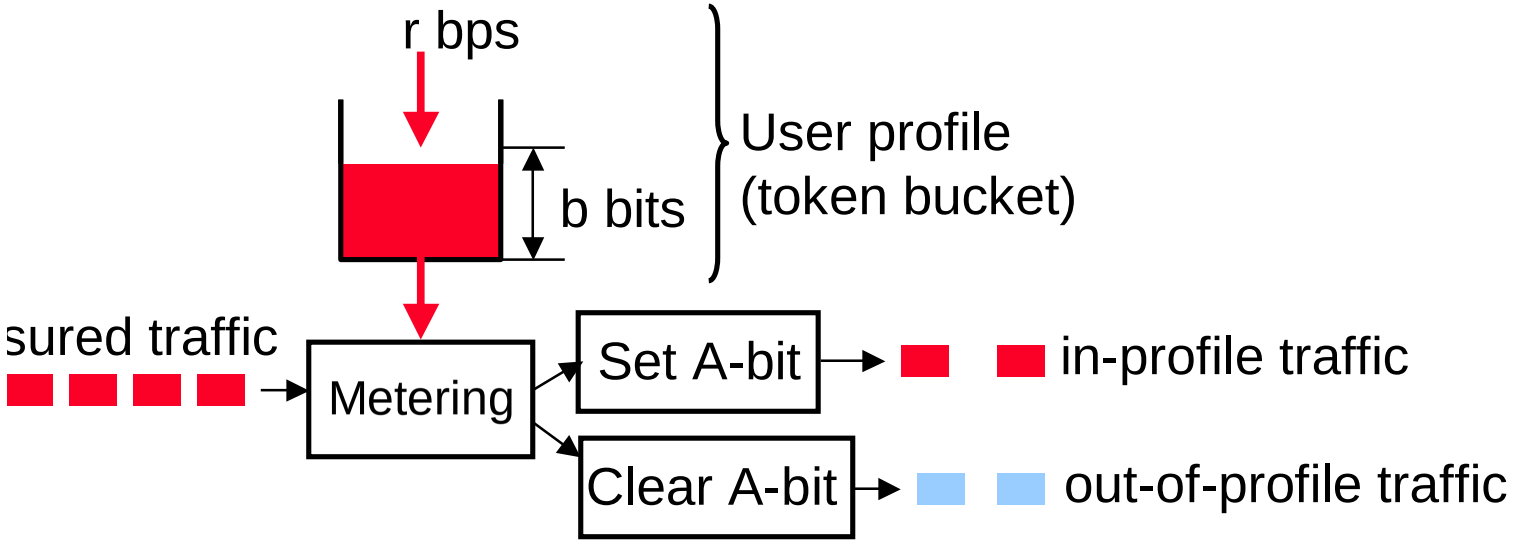

Ingress
Traffic conditioner
Class 1
Marked traffic
Traffic conditioner
Class 2
a traffic
Classifier
Scheduler
Best-effort
r aggregate
assification
g., user)

Assumptions

- Assume two bits
  - P-bit denotes premium traffic
  - A-bit denotes assured traffic
- Traffic conditioner (TC) implement
  - Metering
  - Marking
  - Shaping

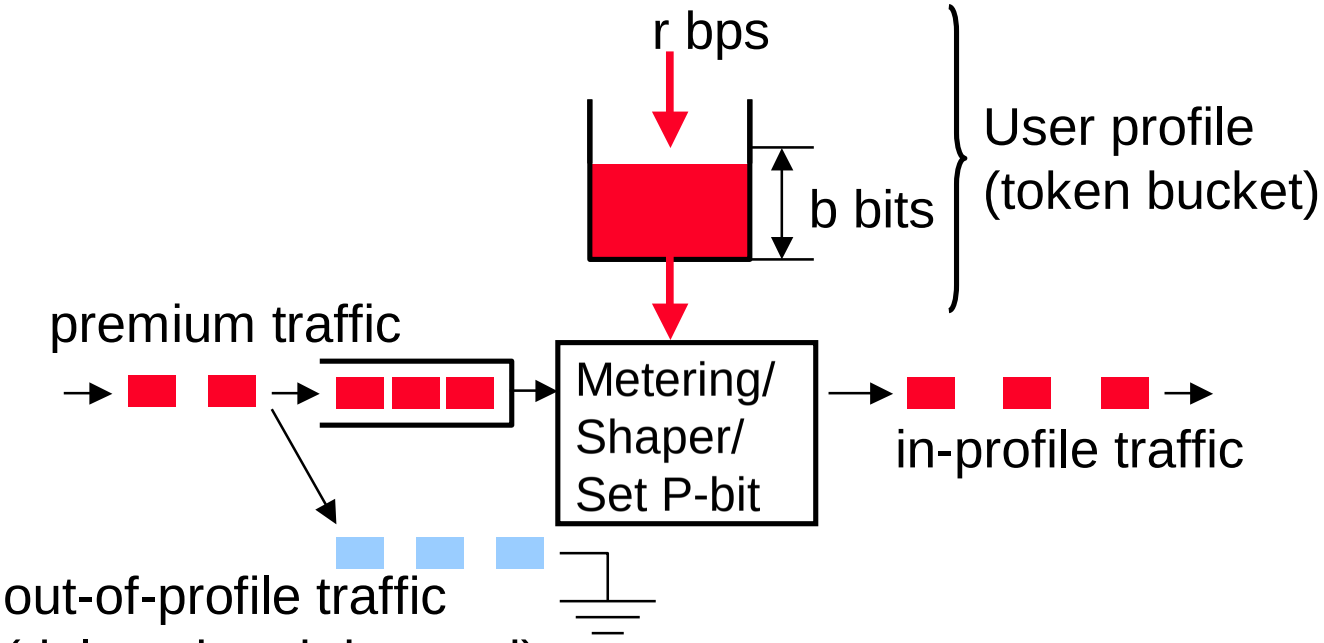
TC Performing Metering/Marking

- Used to implement Assured Service
- In-profile traffic is marked:
  - A-bit is set in every packet
- Out-of-profile (excess) traffic is unmarked
  - A-bit is cleared (if it was previously set) in every packet; this traffic treated as best-effort



TC Performing Metering/Marking/Shaping

- Used to implement Premium Service
- In-profile traffic marked:
  - Set P-bit in each packet
- Out-of-profile traffic is delayed, and when buffer overflows it is dropped



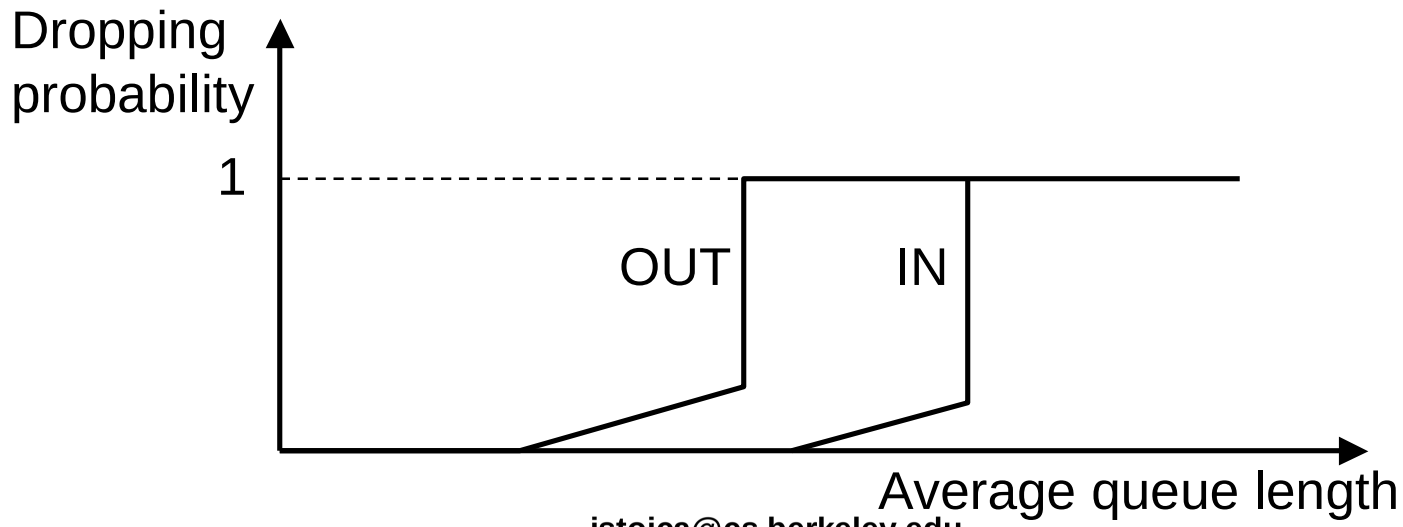
Scheduler

Employed by both edge and core routers

For premium service – use strict priority, or weighted fair queuing (WFQ)

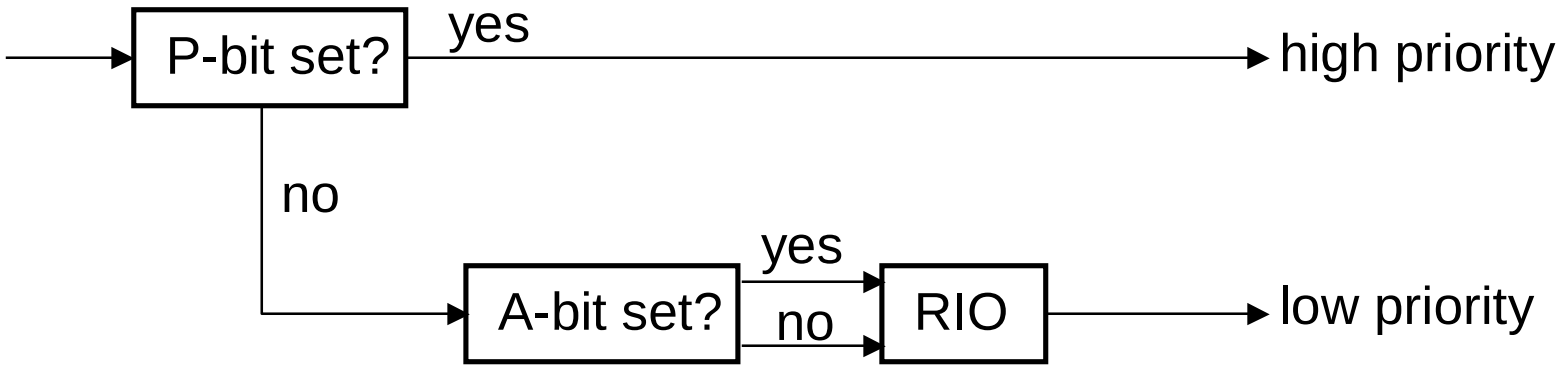
For assured service – use RIO (RED with In and Out)

- Always drop OUT packets first
  - For OUT measure entire queue
  - For IN measure only in-profile queue



Scheduler Example

- Premium traffic sent at high priority
- Assured and best-effort traffic pass through RIO and then sent at low priority

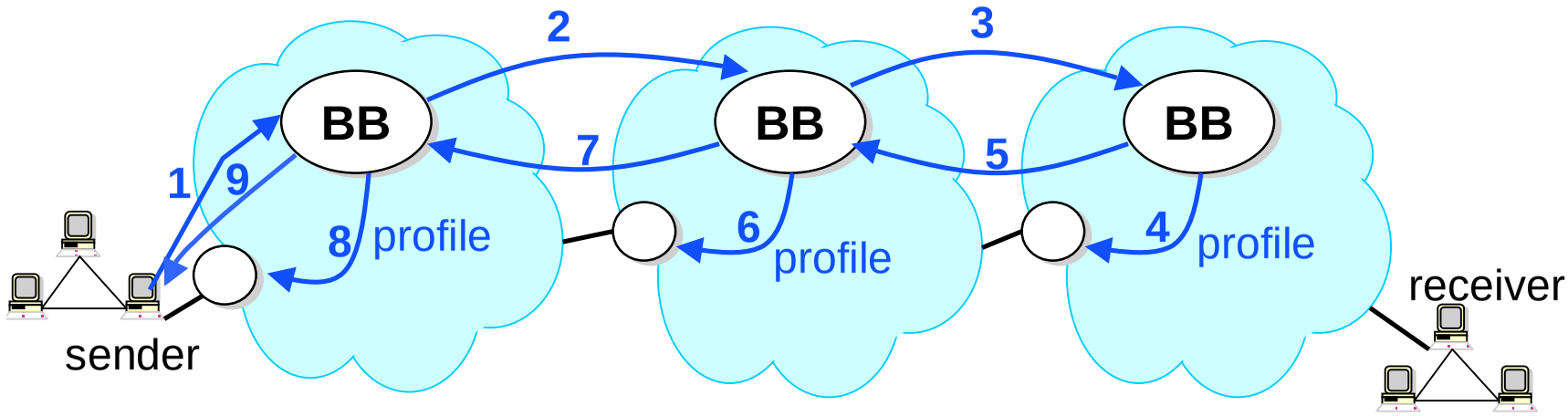


Control Path

- Each domain is assigned a Bandwidth Broker (BB)
  - Usually, used to perform ingress-egress bandwidth allocation
- BB is responsible to perform admission control in the entire domain
- BB not easy to implement
  - Require complete knowledge about domain
  - Single point of failure, may be performance bottleneck
  - Designing BB still a research problem

Example

- Achieve end-to-end bandwidth guarantee



Comparison to Best-Effort and Intserv

	Best-Effort	Diffserv	Intserv
ervice	Connectivity No isolation No guarantees	Per aggregate isolation Per aggregate guarantee	Per flow isolation Per flow guarantee
ervice ope	End-to-end	Domain	End-to-end
omplexity	No setup	Long term setup	Per flow steup
calability	Highly scalable (nodes maintain only routing state)	Scalable (edge routers maintains per aggregate state; core routers per class state)	Not scalable (each router maintains per flow state)

Summary

- Diffserv more scalable than Intserv
  - Edge routers maintain per aggregate state
  - Core routers maintain state only for a few traffic classes
- But, provides weaker services than Intserv, e.g.,
  - Per aggregate bandwidth guarantees (premium service) vs. per flow bandwidth and delay guarantees
- BB is not an entirely solved problem
  - Single point of failure
  - Handle only long term reservations (hours, days)