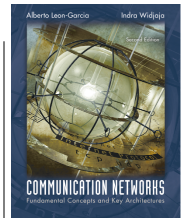
Supplement
Queuing and DiffServ



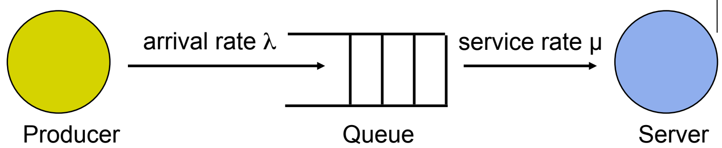
PART I: Preliminaries of Queueing

PART II: Differentiated Services (DiffServ)

Reading Materials

- Chapter 7.7 of the textbook, and two papers as follows:
- *"A Case for Relative Differentiated Services and the Proportional Differentiation Model",* Constantinos Dovrolis and Parameswaran Ramanathan, IEEE Network, Sep/Oct 1999.
- "*Proportional Differentiated Services: Delay Differentiation and Packet Scheduling*", Constantinos Dovrolis*, Dimitrios Stiliadis, and* Parameswaran Ramanathan*, IEEE Transactions on Networking, 10(1), 2002. (preliminary version in ACM SIGCOMM 2000)*

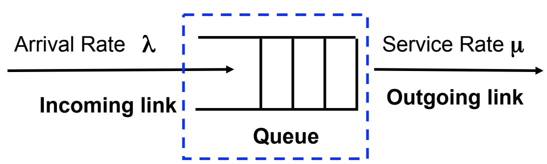
Producer-Server Model



- Throughput:
  - The number of tasks completed by the server in unit time
  - In order to get the highest possible throughput:
    - The server should never be idle
    - The queue should never be empty
- Response time:
  - Begins when a task is placed in the queue
  - Ends when it is completed by the server
  - In order to minimize the response time:
    - The queue should be empty
    - The server will be idle

3

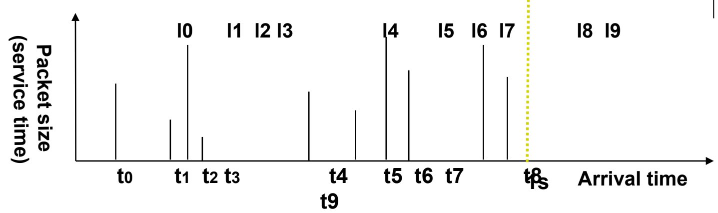
A General Queueing Model in Packet Network



- Arrival rate λ:
  - The number of packets coming from the incoming link in a unit time
- Service rate μ:
  - The number of packets forwarded by the outgoing link in a unit time
- Utilization $\rho = \lambda / \mu$
  - If $\rho > 1$, the queue's length is to be infinite, so is the queueing delay
  - If $\rho = 1$, the queue's length is to be infinite, so is the queueing delay if the inter-arrival distributions are not deterministic; otherwise, finite queue's length, D/D/1 queue
  - If $\rho < 1$, finite queue's length, finite queueing delay!
    - Most scenarios belong to this case!

4

Preliminaries



- Three Components
  - Inter-arrival time distributions (t0, t1-t0, t2-t1, t3-t2, …, t9-t8)
  - Packet length (service time) distributions (l0, l1, l2, l3, …, l9)
  - The number of servers (FIFO queue discipline vs. PS)
- Distributions
  - Deterministic distribution (D)
  - Exponential distribution (M): (ce^{-cx}) – Poisson distribution
  - General distribution (like pareto…)
- Response time (total time) = queueing delay + service time

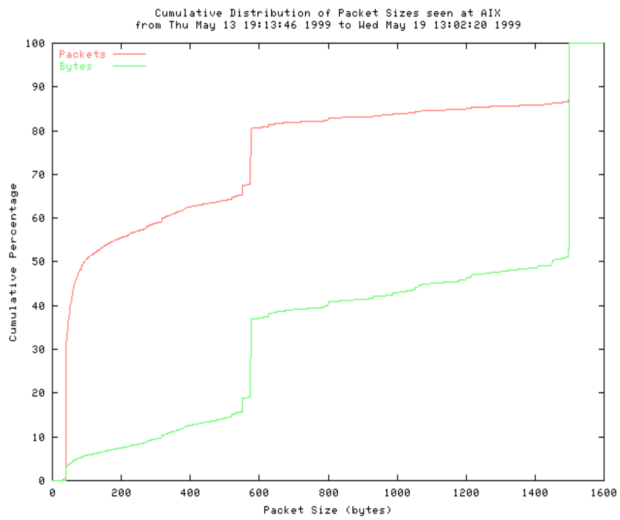
5

Packet Lengths & Service Times

- *R* bits per second transmission rate
- *L* = # bits in a packet
- *X* = *L/R* = time to transmit ("service") a packet
- Packet lengths are usually variable
  - Distribution of lengths → Dist. of service times
  - Common models:
    - Constant packet length (all the same)
    - Exponential distribution
    - Internet Measured Distributions fairly constant
      - See next chart

6

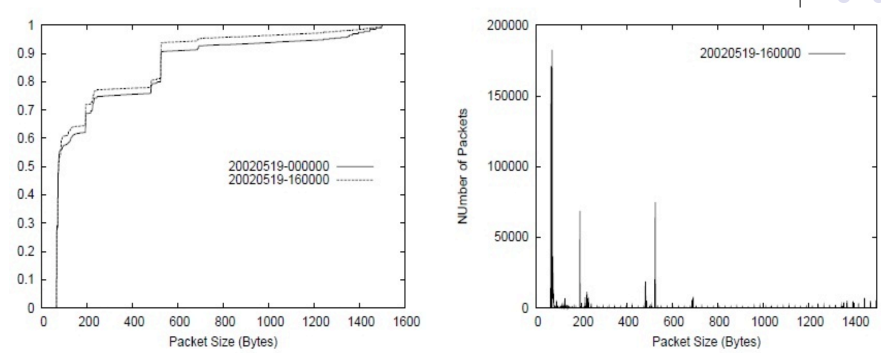
Measure Internet Packet Distribution



- Dominated by TCP traffic (85%)
- ~40% packets are minimum-sized 40 byte packets for TCP ACKs
- ~15% packets are maximum-sized Ethernet 1500 frames
- ~15% packets are 552 & 576 byte packets for TCP implementations that do not use path MTU discovery
- Mean=413 bytes
- Stand Dev=509 bytes
- Source: caida.org

7

IP Packet Distribution – Our Findings



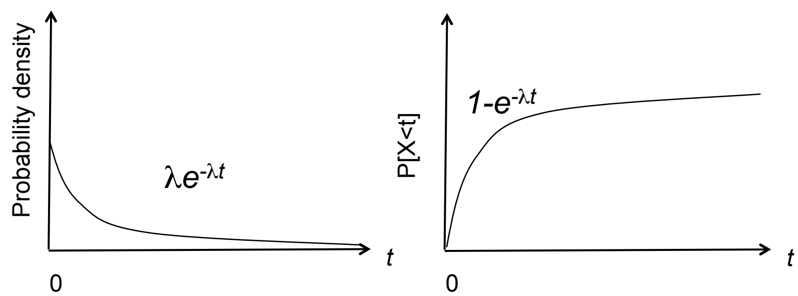
a) The cumulative distribution function. (b) Number of packets for one trace.

Figure : Packet size distributions of two Bell Labs-I IP traces.

8

Poisson Arrivals

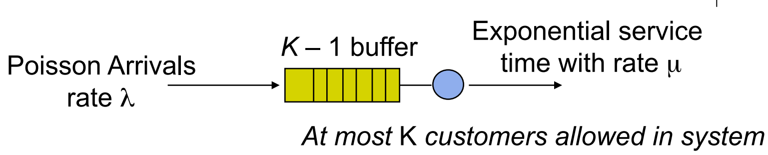
- Average Arrival Rate: λ packets per second
- Arrivals are equally-likely to occur at any point in time
- Time between consecutive arrivals is an exponential random variable with mean 1/ λ



$$P[X > t] = e^{-t/E[X]} = e^{-\lambda t} \quad \text{for } t > 0.$$

9

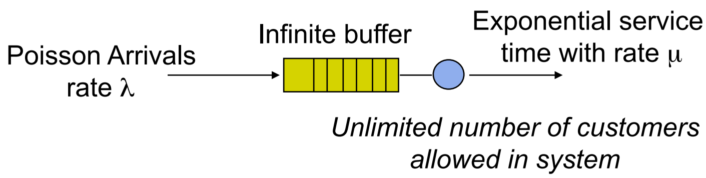
M/M/1/K Queueing Model



- 1 customer served at a time; up to $K - 1$ can wait in queue
- Mean service time E[X] = $1/\mu$
- Key parameter Load: $\rho = \lambda/\mu$
- When $\lambda << \mu$ ($\rho \approx 0$), customers arrive infrequently and usually find system empty, so delay is low and loss is unlikely
- As λ approaches μ ($\rho \to 1$), customers start bunching up and delays increase and losses occur more frequently
- When $\lambda > \mu$ ($\rho > 0$), customers arrive faster than they can be processed, so most customers find system full and those that do enter have to wait about $K - 1$ service times

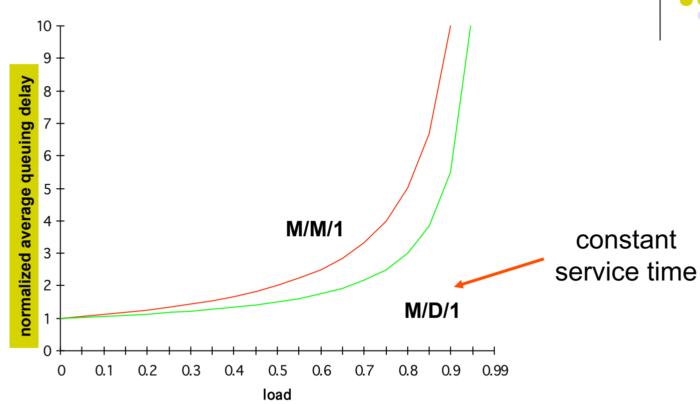
10

M/M/1 Queue



- P_b=0 since customers are never blocked
- Average Time in system E[T] = E[W] + E[X]
- When λ << μ, customers arrive infrequently and delays are low
- As λ approaches μ, customers start bunching up and average delays increase
- When λ > μ, customers arrive faster than they can be processed and queue grows without bound (unstable)

Average Delay in M/M/1 & M/D/1



$$\text{M/M/1}: E[W] = \left[\frac{\rho}{1-\rho}\right]\frac{1}{\mu}; E[T_M] = \left[\frac{\rho}{1-\rho}\right]\frac{1}{\mu} + \frac{1}{\mu} = \frac{1}{\mu}\left[\frac{1}{1-\rho}\right] = \left[\frac{1}{\mu-\lambda}\right]$$

$$\text{M/D/1}: E[T_D] = \left[1 + \frac{\rho}{2(1-\rho)}\right]\frac{1}{\mu} = \left[\frac{\rho}{2(1-\rho)}\right]\frac{1}{\mu} + \frac{1}{\mu}$$

Effect of Scale

- C = 100,000 bps
- Exp. Dist. with Avg. Packet Length: 10,000 bits
- Service Time: X=0.1 second
- Arrival Rate: 7.5 pkts/sec
- Load: ρ=0.75
- Mean Queuing Delay:

$E[W]$ = 0.75/(1 -.75) = 0.3 sec

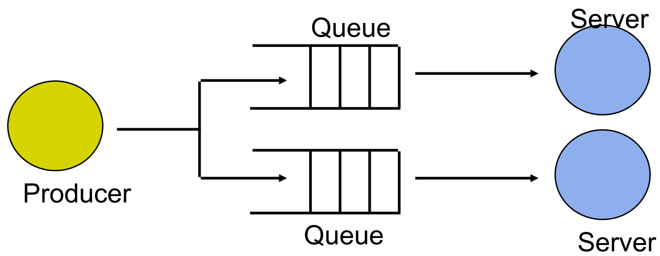
- C = 10,000,000 bps
- Exp. Dist. with Avg. Packet Length: 10,000 bits
- Service Time: X=0.001 second
- Arrival Rate: 750 pkts/sec
- Load: ρ=0.75
- Mean Queuing Delay:
- $E[T]$ = 0.001/(1-.75) = 0.004 sec

Reduction by factor of 100

Aggregation of flows can improve Delay & Loss Performance

13

Throughput Enhancement



- In general throughput can be improved by:
  - Throwing more hardware at the problem
- Response time is much harder to reduce:
  - Ultimately it is limited by the speed of light
  - You cannot bribe God!

14

Differentiated QoS

- Why same-service-to-all model not sufficient?
- What are current QoS differentiation models in networks
- Proportional differentiated services: delay differentiation and packet scheduling
- Proportional differentiated services: loss rate differentiation and packet dropping

Differentiated Services

- Different users have different service expectations for different services
- Internet Engineering Task Force (IETF), April 1998
- The goal is to define configurable types of packet forwarding (called Per-Hop Behaviors, PHBs), which can provide local (per-hop) service differentiation for large aggregates of network traffic, as opposed to end-to-end performance guarantees for individual flows.
  - Stateless priority mechanisms at the network core + stateful mechanisms at the network edge

Best-effort services
(Same-service-to-all)

Integrated Services
(Reservations-based)

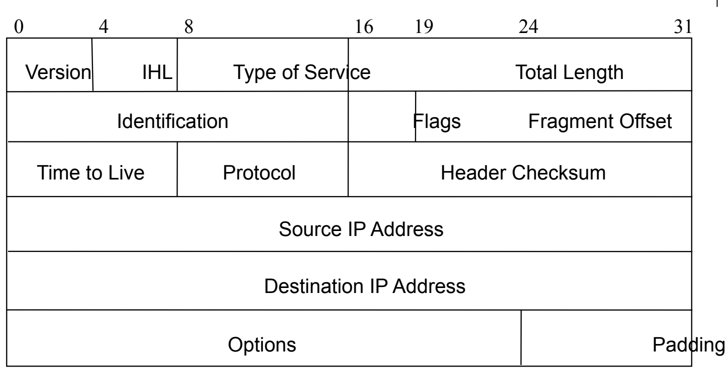
Differentiated Services
(relative vs. absolute)

IntServ vs. DiffServ

	Integrated services	Differentiated services
Granularity of service differentiation	Individual flow	Aggregate of flows
State in routers (e.g., scheduling, buffer management)	Per-flow	Per-aggregate
Traffic classification basis	Several header fields	The DS field (6 bits) of the IP header
Type of service differentiation	Deterministic or statistical guarantees	Absolute or relative assurances
Admission control	Required	Required for absolute differentiation only
Coordination for service differentiation	End-to-end	Local (per-hop)
Scope of service differentiation	A unicast or multicast path	Anywhere in a network or in specific paths
Scalability	Limited by the number of flows	Limited by the number of classes of service
Network accounting	Based on flow characteristics and QoS requirement	Based on class usage
Network management	Similar to circuit-switched networks	Similar to existing IP networks

17

IP Packet Header



Type of service (TOS): traditionally priority of packet at each router. Recent Differentiated Services redefines TOS field to include other services besides best effort.

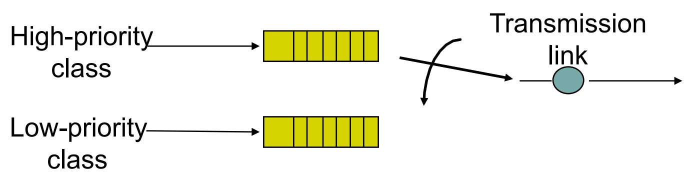
18

Properties and Models

- Properties: Predictability, Controllability, Fairness
- Characteristics: bandwidth, delay, jitter, loss
- Absolute differentiated services
  - Clients receive an absolute service profile from networks
- Relative differentiated services (network traffic in *N* classes)
  - Higher classes will be relatively better (or no worse) than lower classes;
  - Strict Prioritization: delay aspect and loss aspect
  - Price Differentiation: assumption of higher prices leading to lower loads and hence better services
  - Capacity Differentiation: allocates forwarding resources (bandwidth and buffer) between classes
  - Proportional Differentiation: independent of class load

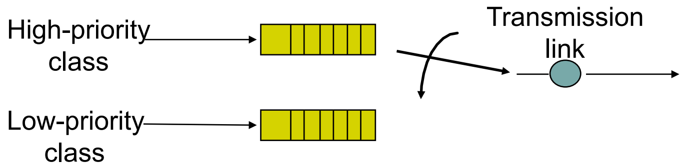
Proportional Delay Differentiation

- Consider a lossless, work-conserving, and non-preemptive link with capacity C (bytes per second) services N first come first served (FCFS) queues, one for each traffic class.
  - Assume that *N* classes have the same packet size distribution.
  - P1: No isolation between classes, one class increases workload, the delays of all classes will also increase
  - P2: increasing the rate of a higher class causes a larger increase in the average class delays than increasing the rate of a lower class



PDD Model

- The model states that certain class performance metrics should be proportional to the differentiation parameters that the network operator chooses.



$$\frac{\bar{q}_i(t, t+\tau)}{\bar{q}_j(t, t+\tau)} = \frac{c_i}{c_j}$$

Question: how to honor the proportionality in terms of average packet queueing delay and loss rate?

21

Packet Average Delay Scheduler (PAD)

- Proportional delay differentiation: normalized average delays be equal in all classes

$$\tilde{d}_i = \frac{\bar{d}_i}{\delta_i} = \frac{\bar{d}_j}{\delta_j} = \tilde{d}_j \qquad 1 \le i, j \le N$$

  - The normalized average delay of class at *t* is (assuming backlogged at *t*)

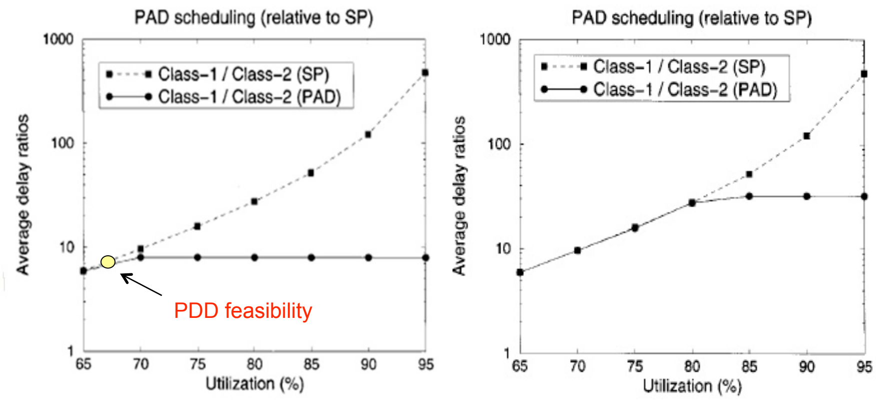
$$\tilde{d}_i(t) = \frac{1}{\delta_i} \frac{\sum_{m=1}^{|D_i(t)|} d_i^m}{|D_i(t)|} = \frac{1}{\delta_i} \frac{S_i}{P_i}$$

- PAD chooses the backlogged class *j* with the maximum normalized average delay

$$j = \arg\max_{i \in B(t)} \tilde{d}_i(t).$$

22

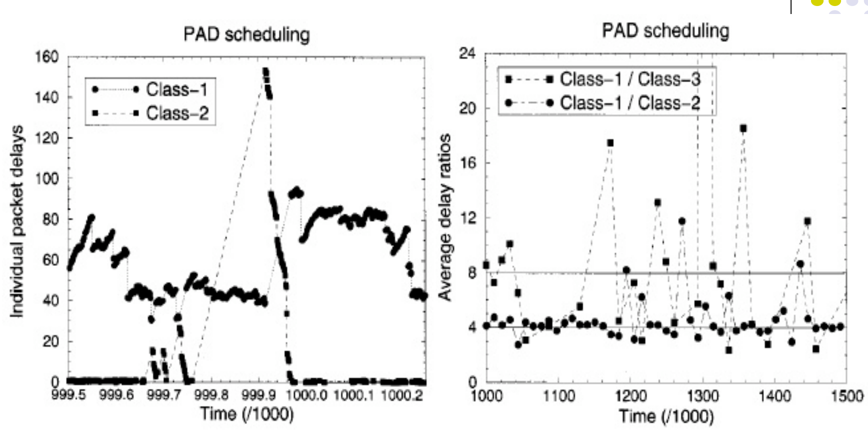
PAD Long-term Performance (vs. Strict Priority)



(a) Delay ratio is 1:8. (b) Delay ratio is 1:32.

(a uniform load distribution λ1 = λ2)

PAD Short-term Performance



(a) Individual packets. (b) Short-term delay ratios.

(μ = 90%; load distribution λ1 : λ2 = 7 : 3)

What did you observe? Good or poor delay differentiation in short-scale? Predictable or unpredictable delay differentiation (any violation)? Why?

Waiting Time Priority Scheduler (WTP)

- Proportional delay differentiation: normalized head waiting time be equal in all classes (inherited from Time-dependent priorities: priority increases proportionally to the packet's waiting time)

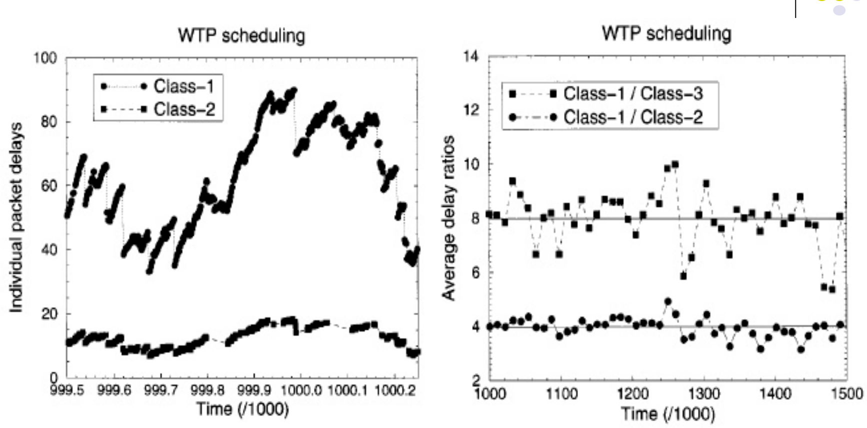
$$\tilde{w}_i(t) = w_i(t)/\delta_i \qquad 1 \le i \le N$$

  - The normalized average delay of class at *t* is

$$j = \arg\max_{i \in B(t)} \tilde{w}_i(t).$$

  - WTP chooses the backlogged class *j* with the maximum normalized head waiting time

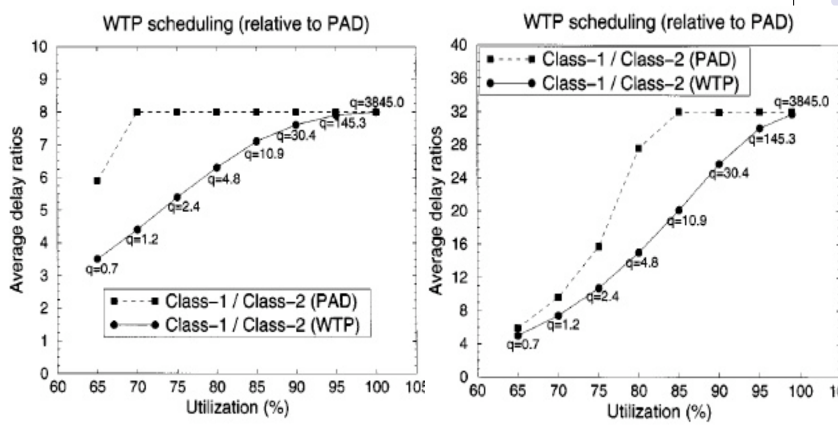
WTP Short-term Performance (vs. PAD)



(a) Individual packets.

(b) Short-term delay ratios.

WTP Long-term Performance (vs. PAD)



(a) Delay ratio is 1:8. (b) Delay ratio is 1:32.

What did you observe? How to compromise long-term and short-term perf.?

Hybrid Packet Delay Scheduler (HPD)

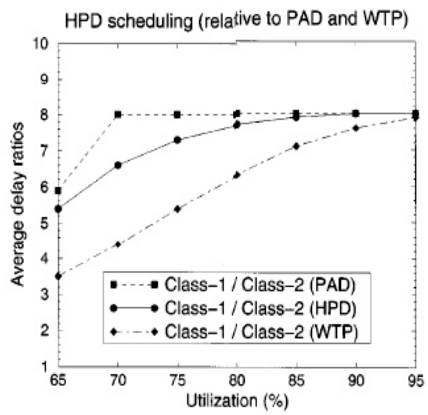
- **Proportional delay differentiation: normalized hybrid delay be equal in all classes**
  - **The normalized hybrid packet delay of class at *t* is**

$$\tilde{h}_i(t) = g \cdot \tilde{d}_i(t) + (1-g) \cdot \tilde{w}_i(t).$$

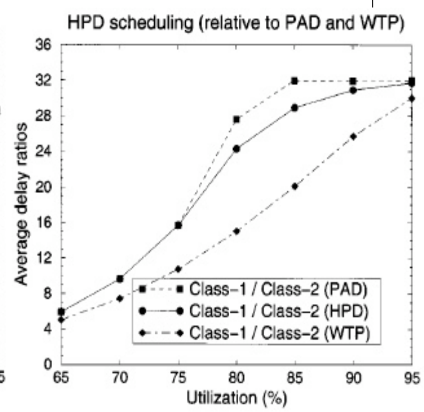
- **WTP chooses the backlogged class *j* with the maximum normalized head waiting time**

$$j = \arg\max_{i \in B(t)} \tilde{h}_i(t).$$

HPD Long-term Performance



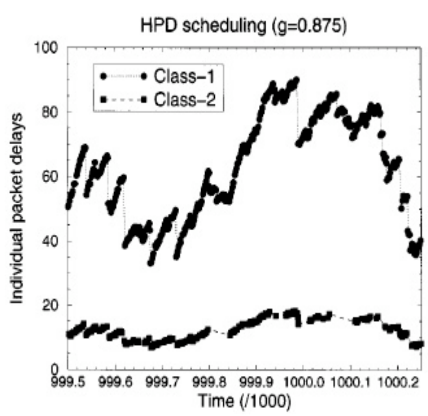
(a) Delay ratio is 1:8.



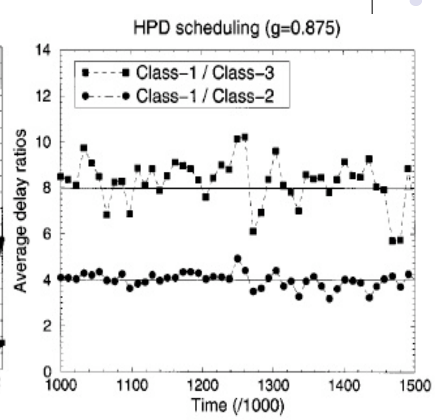
(b) Delay ratio is 1:32.

How to compare HPD with WTP for the short-term perferformance?

HPD Short-term Performance



(a) Individual packets.

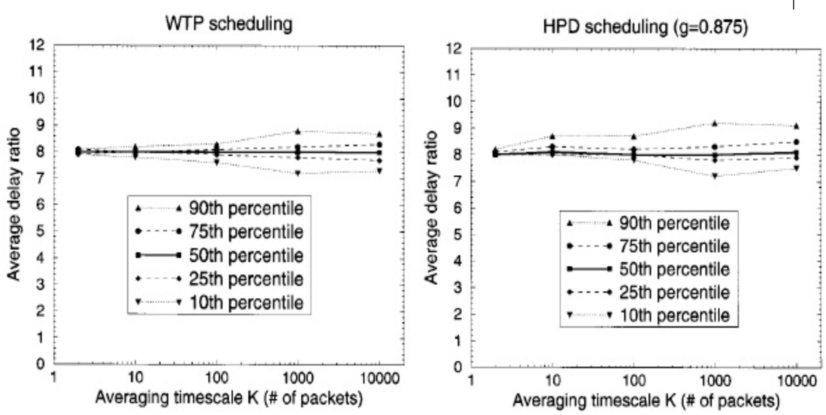


(b) Short-term delay ratios.

When g = 0. 875, HPD provides almost indistinguishable results with WTP.

How to compare the performance with that of WTP more quantitatively?

HPD Short-term Performance II

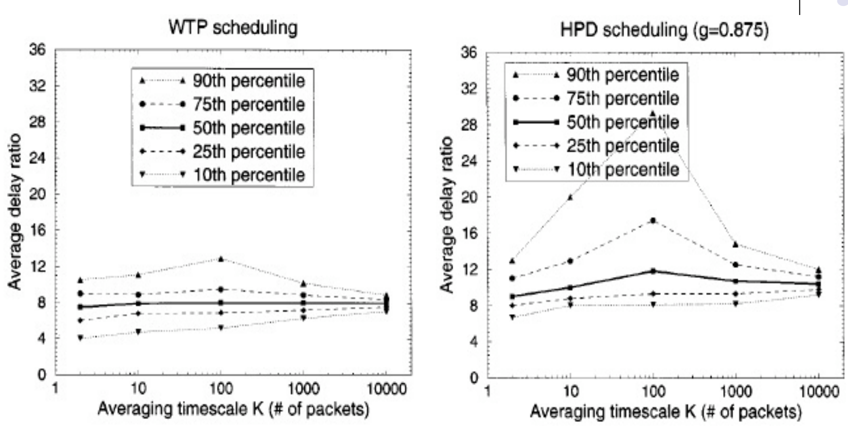


Under heavy load condition (utilization = **95%**).

What is your observation in terms of the short-term performance comparison?

31

HPD Short-term Performance III



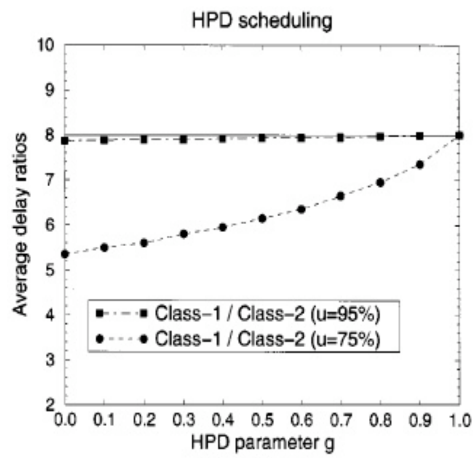
Under moderate load condition (utilization = **80%**).

What is your observation in terms of the short-term performance comparison?

Any other sensitivity study you can think of? How to choose a good *g*?

32

HPD: Impact of the parameter *g*



How about the PDD feasibility regarding to the workload distributions?

33

Reading and Homework

- Chapter 5
- Reading: papers # 3 & #4; see course Web site
  - Due Oct 10, Monday
- Homework (due Friday, 5:00PM)
  - 5.3, 5.4(a)&(b), 5.7, 5.19, 5.22, 5.52, 5.53