# Characterization and Performance Evaluation for the Proportional Delay Differentiated Services

Matthew Leung, John C.S. Lui
The Chinese University of Hong Kong

David Yau
Purdue University

# OUTLINE

- Differentiated Services
- Waiting Time Priority (WTP) Scheduling
  - performance property
  - background
  - feasibility region (necessary and/or sufficient conditions)
  - determination of *control parameters*
- Experiments
- Conclusion

# Introduction: What is DS?

- To provide different treatment to different classes of packet.
- IP packets are classified into classes.
- Classes are recognized by the *tag* in the IP header.
- **Edge router:** responsible for classifying packets into classes.
- **Core router:** uses some specially designed packet scheduling and buffer management algorithms to provide different *per-hop behavior* (PHB) so different classes will have different performance (e.g, packet delay, throughput,..etc).
- Need to understand and analyze these packet scheduling algorithms so as to understand the performance and implication.

# Proportional Delay Differentiated Service

- Why *pay* higher price unless the ISP can provide *consistent* differentiation?
- Example: FTP.
- One possible approach is to provide *proportional delay*.
- Let $W_i$ be the expected waiting time of class $i$ traffic. We want to achieve

$$\frac{W_p}{W_{p+1}} = r_{p,p+1}$$

- For example: $\frac{W_1}{W_2} = 1.5$ implies the waiting time of class 1 traffic is 50% more than class 2 traffic.

# Waiting-Time Priority (WTP)

- C. Dovrolis et.al in SIGCOMM'99 proposed two algorithms to realize the proportional delay differentiated service.
  - The Backlog Proportional Rate (BPR) scheduler
  - The Waiting Time Priority (WTP) scheduler
  - Performance evaluation via simulation and the result shows that WTP is better than BPR.
- Problems:
  - Proportional waiting can only be guaranteed at *high utilization*.
  - **Reason**: the control parameters of WTP are not chosen correctly.

# Our contribution

- Mathematical analysis of WTP.
- Provide understanding and intuition to WTP (e.g. control parameters depend on system utilization).
- Determine the *appropriate value* for the control parameters.
- Determine the *feasible region* of our control problem.

# Time Dependent Priority (TDP)

- WTP is equivalent to the time dependent priority (TDP) scheduler.
- TDP is a non-preemptive scheduler.
- When a packet from class $p$ arrives at time $\tau$, its priority at time $t$, for $t > \tau$, is $(t-\tau)b_p$, where $b_p$ is the **control parameter** for class $p$.
- Server selects the highest *instantaneous* priority packet for transmission.

# TDP: continue

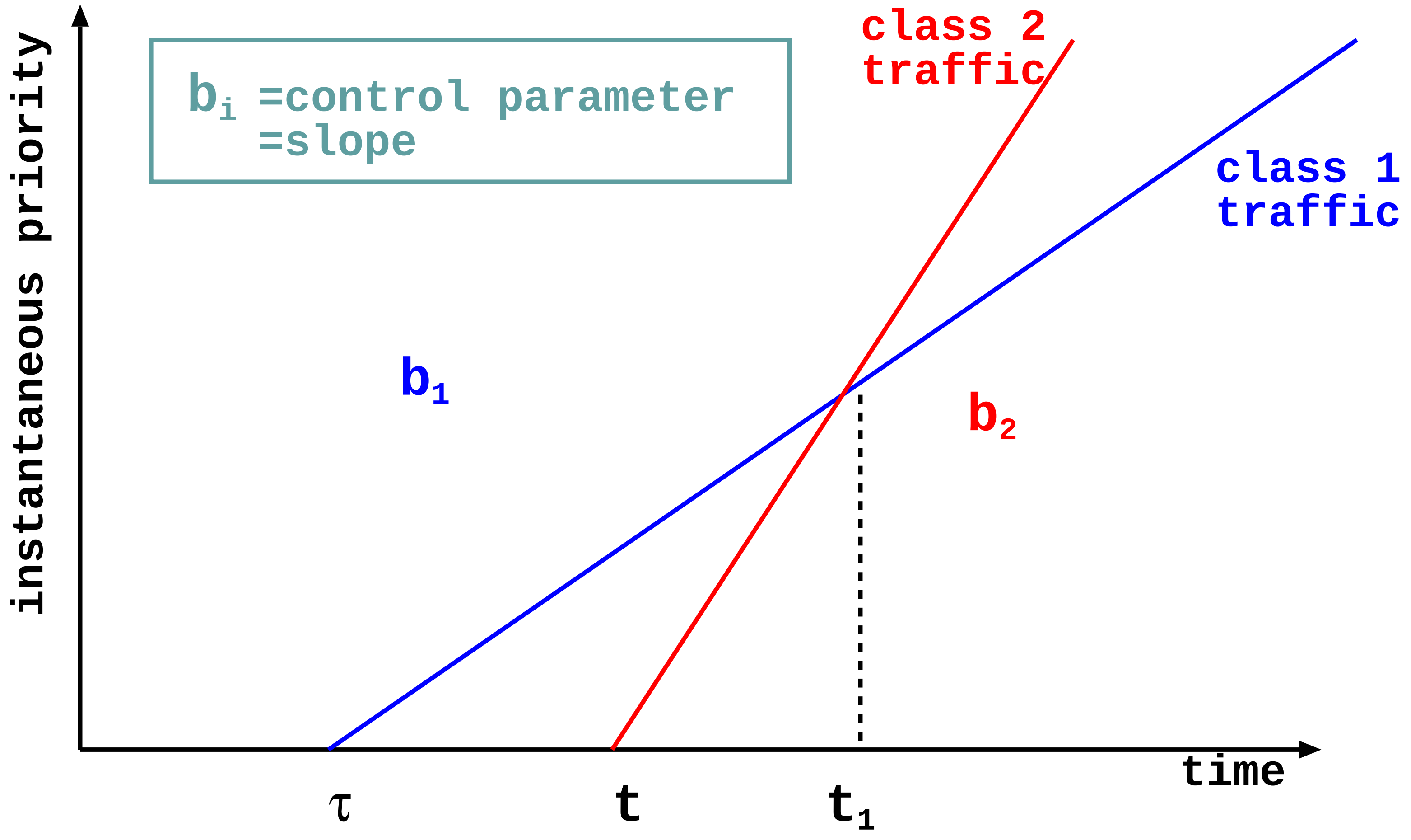


Figure 1: An illustration of the TDP scheduler

# WTP: Performance Analysis

- Assume Poisson arrival ($\lambda_i$), general service time ($\overline{x_i}, \overline{x_i^2}$) and $N > 1$ classes of traffic.

- We have the closed form expression for $W_p$, the average waiting time for class $p$ traffic, which is:

$$W_p \;=\; \frac{[W_0/(1-\rho)] - \sum_{i=1}^{p-1} \rho_i W_i \left[1-(b_i/b_p)\right]}{1-\sum_{i=p+1}^{N} \rho_i \left[1-(b_p/b_i)\right]} \qquad p = 1, 2, \ldots, N$$

where $\quad W_0 = \dfrac{\sum_{i=1}^{N} \lambda_i \overline{x_i^2}}{2}$

## Simple Case: $N = 2$

**Theorem 1:** *Let $r_{1,2}$ be the specified ratio of the average waiting time of class 1 packets to that of class 2 packets. To achieve the specified performance ration $r_{1,2}$, the* **necessary and sufficient** *condition is that the system utilization*

$$\rho > 1 - \frac{1}{r_{1,2}}$$

**Corollary:** *If the necessary and sufficient condition is satisfied, then to achieve $W_1/W_2 = r_{1,2}$, we set the control parameters:*

$$b_1 = 1 \quad ; \quad b_2 = \frac{\rho}{\rho - 1 + 1/r_{1,2}}$$

## Necessary feasibility condition for $N > 2$

- Put all $b_p$'s to one side,

$$b_p\left(W_p\sum_{i=p+1}^{N}\frac{\rho_i}{b_i}\right)-\frac{1}{b_p}\left(\sum_{i=1}^{p-1}\rho_i W_i b_i\right)=$$

$$\frac{W_0}{1-\rho}-\sum_{i=1}^{p-1}\rho_i W_i-\left(1-\sum_{i=p+1}^{N}\rho_i\right)W_p$$

- After rewriting, we have:

$$b_p F(p)-\frac{G(p)}{b_p}=R(p)\qquad p=1,2,\ldots,N$$

## Continue: For $N > 2$

**Theorem 2:** *The necessary conditions to have positive solution for the control parameters $b_p$'s are that $R(1) > 0$ and $R(N) < 0$.*

**Implication :** $R(1) > 0$ implies

$$\frac{W_0}{W_1(1-\rho)} > \left(1 - \sum_{i=2}^{N} \rho_i\right).$$

- If we want a large waiting time differentiation, $W_1$ has to be large (since $W_1 \geq W_2 \geq \cdots W_n$). Based on *conservation law*, this implies that the fraction on the L.H.S. has to be small, thus $\sum_{i=2}^{N} \rho_i$ should be close to one !
- The *physical meaning* is that there should be enough packets from higher classes to keep the server busy so that the lower classes packets are delayed.

## Continue: For $N > 2$

**Second Implication :** $R(N) < 0$ implies

$$\frac{W_0}{1-\rho} - \sum_{i=1}^{N-1} \rho_i W_i - W_N < 0.$$

- To realize the inequality, the magnitude of the negative part should be large. So $\rho_i$ should be large, especially the lower classes.
- The *physical meaning* is that the server has to delay the packets from lower classes. If their loading is high, many of them will be backlogged and their waiting time will increase.

**Last implication:** even when $\rho$ is fixed, realization depends on the *distribution* of $\rho_i$.

# Computation Algorithm

- Solve the system of non-linear equations by Gauss-Seidel iteration method.
- Change the equations into the form:

$$b_p = \phi_p(b_1, b_2, \ldots, b_N) \qquad\qquad p = 1, 2, \ldots, N$$

where

$$\phi_i(b_1, b_2, \ldots, b_N) = \begin{cases} \frac{R(p)+B(p)/R(p)}{A(p)} & \text{for } p \neq N \\ \\ \sqrt{\frac{b_p B(p)}{-R(N)}} & \text{for } p = N. \end{cases}$$

# Computation Algorithm

---

**Procedure: Iterative Algorithm**

**Input:** $\lambda_i$, $\overline{x_i}$, $\overline{x_i^2}$ for $i = 1, .., N$.

/*arrival rates, $1^{st}$, $2^{nd}$ moments of service times*/

**Output:** $\boldsymbol{b} = [b_1, b_2, \ldots, b_N]$.

**begin**

    $k = 0$ and $b_p^{(0)} = \frac{1}{W_p}$ for $p = 1, \ldots, N$; /* initialize */

    /* test for convergence */

    **while** $((\sum_{p=1}^{N} |f_p(\boldsymbol{b}^{(k)}| > \epsilon)$ and $k <$ MAX_ITERATION_COUNT )

        **begin**

            **for** $(p = 1; p <= N; p = p + 1)$

                $b_p^{(k+1)} = \phi_p(b_1^{(k+1)}, b_2^{(k+1)}, \ldots, b_{p-1}^{(k+1)}, b_p^{(k)}, \ldots, b_N^{(k)})$;

            $k = k + 1$;

        **end**

**end**

---

## Convergence Criteria

- The iteration converges iff the Jacobi matrix of $\phi_p$ has all its eigenvalues less than 1 in moduli
- Since the eigenvalues are not easily obtainable, there is a sufficient condition:
  - Row-sum or column sum criterion.

# Experiments

- Poisson arrivals
- Packet size distribution:
  - 40% packets has 40 bytes
  - 50% packets has 550 bytes
  - 10% packets has 1500 bytes
- Output link capacity is 441 bytes
- mean service time = 1.0

## Medium Utilization ($\rho = 0.6$)

- Total of 3 classes of packets
- Target waiting time spacing is 1.5.
  - $\frac{W_1}{W_2} = 1.5$ and $\frac{W_2}{W_3} = 1.5$.
- Mean arrival rates are 0.2 for all classes.
- Using our iterative algorithm, we have $\boldsymbol{b} = [1, 2.6, 7.4]$. We achieve:
  - $\frac{W_1}{W_2} = 1.54$ and $\frac{W_2}{W_3} = 1.48$.
- Using Dovrolis (SIGCOMM'99) result, we set $\boldsymbol{b} = [1, 1.5, 2.25]$. We achieve:
  - $\frac{W_1}{W_2} = 1.25$ and $\frac{W_2}{W_3} = 1.22$.

## High Utilization ($\rho = 0.95$)

- Total of 3 classes of packets
- Target waiting time spacing is 4.0.
  - $\frac{W_1}{W_2} = 4.0$ and $\frac{W_2}{W_3} = 4.0$.
- Mean arrival rates are: $0.35, 0.3, 0.3$ respectively.
- Using our iterative algorithm, we have $\boldsymbol{b} = [1, 5.1, 35.9]$. We achieve:
  - $\frac{W_1}{W_2} = 3.99$ and $\frac{W_2}{W_3} = 3.89$.
- Using Dovrolis (SIGCOMM'99) result, we set $\boldsymbol{b} = [1, 4, 16]$. We achieve:
  - $\frac{W_1}{W_2} = 3.37$ and $\frac{W_2}{W_3} = 3.03$.

## Performance on short-time scale

- Measure the average waiting time of packets that get served within a specified *time window*.
- The length of the window size includes $100, 500, 1000, 3000, 7000$ and $1000$ p-units, where a p-unit is the average packet transmission time.
- Now the measure is *not* average waiting time, but rather, distribution of waiting time around the target waiting time ratio.
- Three classes, $r_{i,i+1} = 1.1$, $\rho = 0.6$ and $0.9$.

for picture

for picture

## Observation

- When the system utilization is high, a lot of packets are served even in a very short time period.
- When the utilization is low, it takes a longer time for a sufficient number of packets to be served so that the sample mean of the waiting time can reach their actual mean (Law of large number).

# Conclusion

- Want service differentiation for performance and pricing.
- We consider the WTP scheduler to achieve delay proportional differentiated service.
- For two classes, obtain necessary and sufficient condition to achieve a given delay spacing.
- For multiple classes, obtain necessary conditions to test feasibility region.
- Proposed efficient iterative method to compute the values of the control parameters.
- **We provide a more formal approach to gain some understanding of the system.**

## Extension

- Sensitivity and applicability to non-Poisson arrival processes.
- Real-time monitoring and measurement of system parameters.
- Dynamic adjustment of system control parameters so as to adopt to changes in the workload.