Lab

4

Bounded Delay Service

Teams: This lab may be completed in teams of 2 students (Teams of three or more are not permitted. All members receive the same grade).

Purpose of this lab:

In this lab you design traffic characterization and admission control test for a network node that supports a bounded delay service, that is, a service that guarantees that traffic from flows does not exceed given delay bounds.

Software Tools:

- This lab does not require transmission of traffic as in Labs 2 and 3.
- The data analysis and computations can be completed with Matlab, Java, or C/C++.

What to turn in:

- Turn in a hard copy of the lab report, including the plots, printouts of your code, and the anonymous feedback form.

Version 1 (March 20, 2007)

Table of Content

Background on Bounded Delay Service and Admission Control

A bounded delay is a network service that provides guarantees on the maximum delay of traffic. A network that offers a bounded delay service requires additional components.

Traffic specification, service guarantees: Before a new flow with a bounded delay can start, an application must submit a request for this flow, which includes a traffic specification and the desired delay guarantees. Typically, an application specifies traffic in terms of the parameters of a leaky bucket (average rate r, burst size B), or a peak-rate constrained leaky bucket (peak rate P, average rate r, burst size B). Service guarantees are expressed in terms of delay bounds (D^*).

Admission control test: Before a network accepts a new flow with delay guarantees, it performs admission control tests that determine if the network can support the requested service guarantees for the new flow without violating service guarantees made to other flows in the network. If the guarantees of the new flow can be supported, the network admits the flow. Otherwise, the request for the new flow is rejected.

Traffic enforcement: After a flow is admitted, network ensures that the traffic from the flow does not exceed its traffic specification. This is done by passing traffic that enters the network through a traffic regulator (e.g., a leaky bucket). When the traffic exceeds the traffic specification, the excess traffic is either delayed, discarded or marked with a lower priority.

In this lab you explore the design a network service with deterministically bounded delays, called a **bounded delay service**. A design goal for a network with a bounded delay service is to utilize its resources well. The degree to which a network with a bounded delay service can utilize its resources is largely determined by the design of:

- The method used for the characterization of network traffic.
- The scheduling algorithms at the output links of packet switches.
- The quality (accuracy) of the admission control tests.

Part 1. Traffic specification of video traffic

This part is similar to Part 4 of Lab 2, in that it deals with the selection of leaky bucket parameters. The difference is that, here, you only determine the parameters, but do not implement the leaky bucket.

The objective is to obtain a traffic specification for a given VBR video sequence. A video sequence consists of a sequence of frames with (more or less) regular time intervals between frames.

If the sequence of a frame sizes of video source j is given by f_1, f_2, …, f_N and if the time of the *i*-th frame is given by T, we can write an arrival function at the time of the arrival times as:

$$A(n \cdot T) = \sum_{k=1}^{n} f_k$$

A traffic characterization of a video source for a bounded delay services requires a traffic specification that provides an upper bound on the amount of traffic over an interval of time. We refer to such a characterization as a **worst-case traffic specification.**

Conceptually, creating a worst-case traffic specification consists of finding an envelope $E(t)$ for the video trace, that satisfies:

$$E(t) \geq A(t+\tau) - A(\tau), \forall t, \tau$$

or, equivalently, $A(t) \geq E * A(t)$, for all t ≥ 0.

The most amount of traffic generated by a traffic source with arrival function A over a time period of length t can be exactly stated as:

$$E^*(t) = \sup_{\tau \geq 0} \{A(t+\tau) - A(\tau)\}$$

or, equivalently, $E^*(t) = A \oslash A(t)$. The function E^* is called the **empirical envelope.**[1] It is easy to see that it is the best worst-case traffic specification in the sense that no envelope for an arrival function A can be smaller than the empirical envelope.

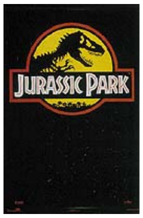
A traffic specification for a video source in terms of an empirical envelope E^* yields a tight envelope, but it requires as many parameters as there are video frame. For a full-length motion picture, this results in 100,000 – 200,000 parameters!

A traffic specification with leaky buckets is a practical alternative. A leaky bucket describes traffic with a small number of parameters. In addition, it is relatively easy to built a traffic regulator that enforces a leaky bucket envelope function (see Lab 2).

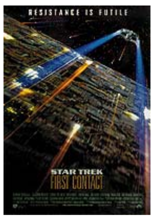
[1] The attribute "empirical" emphasizes that the envelope is derived firsthand from the traffic source, without external assumptions or models.

Exercise 1.1 Empirical envelopes

Consider the following two trace files of video sources available at:

- 

"Jurasic Park": http://trace.eas.asu.edu/TRACE/pics/FrameTrace/h263/Terse_Jurassic256.dat

- 

"Star Trek: First Contact": http://trace.eas.asu.edu/TRACE/pics/FrameTrace/h263/Terse_FirstContact256.dat

The tracefiles have one column indicating the size of a video frame measured in Bytes.

```
Length [byte]
#______________________________________________
     2241
     1877
     3199
     2817
```

Assume that the time interval between video frames is constant at T = 40 ms. The arrival time of the n-th frame is given by n·40 ms.

Arrival time (ms)	Frame size (byte)
40	$f_1 = 2241$
80	$f_2 = 1877$
120	$f_3 = 3199$
160	$f_4 = 2817$
…	…

- In this lab, only the first 1000 lines of each file will be used. You may truncate the file.
- Compute the empirical envelope $E^*(t)$.
  Hint: For the given arrival scenario, $E^*(t)$ is a step function that changes only at the arrival times of new frames, given by:
  $$E^*(n \cdot T) = \max_{0 \le k < N-n} \sum_{j=1}^{n} f_{k+j}, \text{ for } n = 1, 2, \ldots, N$$
- For each video source, prepare a plot that shows the cumulative arrival function A(t), and the empirical envelope $E^*(t)$, evaluated at the frame arrival times, T, 2T, …, N·T.

Exercise 1.2 Comparison with peak rate and mean rate

Let the **peak rate** of a video source be defined as the maximum transmission rate of a frame, measured over the duration of a frame interval. With a constant frame time, this gives:

$$\text{Peak rate} = \frac{1}{T} \max_i \{f_i\}$$

Likewise, the **mean rate** of a video source can be defined as the average transmission rate of a frame, measured over the duration of the complete sequence. With a constant frame time, this gives:

$$\text{Average rate} = \frac{1}{N \cdot T} \sum_{j=1}^{N} f_j$$

- Compute the peak rate and the mean rate of both video sources from Exercise 1.1.
- For each video source, prepare a graph for the time interval [0, N·T] ms (with T=40 ms, N = 1000) that includes the average rate, peak rate, and the empirical envelope computed in Exercise 1.1.
  (To permit a direct comparison of the two graphs, the range and scale of the axes should be identical in both graphs).
- Discuss similarities and differences of the two video sources.

Exercise 1.3 Worst-case traffic characterizations with leaky buckets

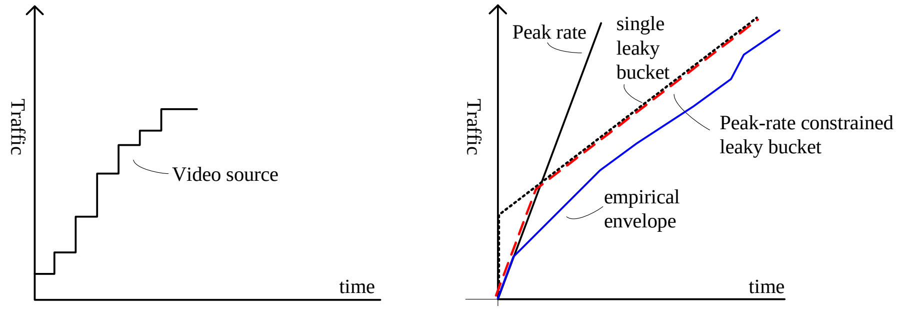
A traffic specification for a video source in terms of an empirical envelope E* yields a tight envelope, but it requires one parameter for each video frame. For a full-length motion picture, this results in 100,000 – 200,000 parameters. Hence, building a traffic regulator that enforces E* is not realistic.

A traffic specification with leaky buckets describes a video source in terms of a sequence of rate and burst parameters. In addition to describing complex with just a few parameters, it is relatively easy to build a traffic regulator for leaky buckets.

In this part of the lab you will obtain a leaky bucket characterization for the *Jurassic Park* and *Star Trek* video sources using the following types of leaky bucket:

Type:	Parameters:	Envelope function:
Leaky bucket	$r > 0, b > 0$	$E(t) = b + r\,t$
Peak-rate constrained leaky bucket	$P > r > 0, b > 0$	$E(t) = \min(P\,t, b + r\,t)$
Dual-Leaky bucket	$P > r > 0, b > M > 0$	$E(t) = \min(M + P\,t, b + r\,t)$

If an empirical envelope E^* for the video source has been determined, a worst-case leaky bucket characterization can be obtained by bounding the empirical envelope from above by a set of linear segments (one for each leaky bucket). This is illustrated in the following figure.



Your task is to make `good' selections for the parameters of the three types of leaky buckets given in the table. The selection of parameters represents a tradeoff between multiple considerations:

1. The average rate allocation should not be much larger than the average rate of the VBR source. Otherwise, there will be an over-allocation of bandwidth.

2. The burstiness of the traffic, as determined by parameters b and M should be small. Otherwise the amount of traffic that can be transmitted into the network at once is too high, resulting in high backlogs and long delays.

For each of the two video sources compute the following:

- For each type of leaky bucket, compute parameters that result in an envelope for the video source. Provide the parameters in a table. Justify the choice of parameters.

- Prepare a plot for the range [0, 1000] ms, that show the three leaky buckets. Also include the empirical envelope (from Exercise 1.2) in the plot.

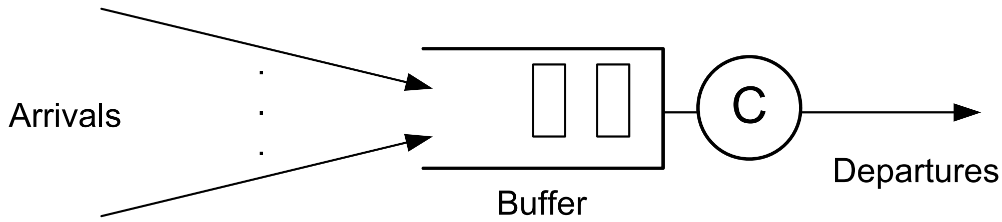
- Discuss the results, taking into considerations the results from Exercise 1.2.

Lab Report:

Include properly labelled and scaled plots and the discussions as requested in the exercises.

Part 2. Admission Control Tests (Homogeneous Traffic)

Consider a single buffered link with capacity C that supports a bounded delay service for video traffic. In this part of the lab, you will run admission tests that determine if the link can support the delay bounds for a given set of flows.



An admission test for a bounded delay service at a link requires a schedulability condition that detects potential violations of delay guarantees at an output link of a packet switch. The schedulability conditions verify that the maximum delay of a packet from any flow does not exceed the specified delay bounds.

In this part of the lab, we consider a scenario where the traffic from all flows consists of the same type of video source, and all flows have the same delay bound. We refer to this scenario as **homogeneous traffic**.

In such a scenario, all scheduling algorithms generate the same FIFO schedule, i.e., traffic from all flows is transmitted in the order of arrival. Let E be an envelope for the video source, let N be the number of flows, and let D^* be the delay bound. Assuming that the traffic at the scheduler consists only of the video data (that is, not accounting for the overhead of packet headers), the schedulability condition is as follows:

$$D^* \geq \sup_{t \geq 0} \left\{ \frac{N \cdot E(t) - Ct}{C} \right\}$$

If the condition holds, no delay bound violation will occur. (Recall that this condition was used in the lecture). Note that it is sufficient to verify this condition at the arrival times of flows, that is, *t = T, 2T, … N T.*

With the schedulability condition above it is feasible to compute the maximum number of flows N_{max} that can be supported with a delaoy bound of D^*. Then, the admission control test simply needs to compare the number of flows on a link. A new flow can be admitted only if the total number of flows (including the new flow) does not exceed N_{max}.

In the following, consider a link with C=100 Mbps. All traffic on the link is either from the *Jurassic Park* or the *Star Trek: First Contact* video source.

Exercise 2.1 Mean rate and peak rate allocation

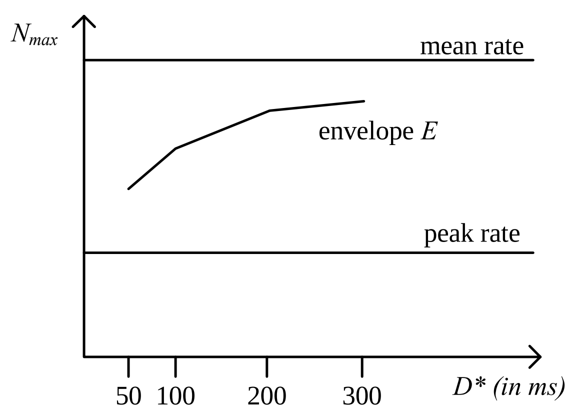
Consider a rate allocation scheme that allocates to each flow a capacity equal to the peak rate and the average rate of the video flows (as defined in Part 1).

- o Determine the maximum number N_{max} of flows of type *Jurassic Park* (and *Star Trek*) that can be allocated on the 100 Mbps link, when each flow is reserved a capacity equal to the peak rate of the flow.
  - o Determine the maximum delay of at the link when the maximum number of flows are allocated.
- o Determine the maximum number N_{max} of flows of type *Jurassic Park* (and *Star Trek*) that can be allocated on the 100 Mbps link, when each flow is reserved a capacity equal to the mean rate of the flow.
  - o Determine the maximum delay of at the link when the maximum number of flows are allocated.

Exercise 2.2 Admission control for homogeneous traffic (Empirical envelope)

Use the schedulability condition for homogenous traffic and compute the maximum number of flows N_{max} that can be allocated on a link with C=100 Mbps as a function of the delay bound D^* of the flow.

In this exercise, the envelope is the empirical envelope of the video source. (Since the empirical envelope E^* is the best possible worst-case characterization, it will result in the largest number of flows N_{max} of flows that can be supported with a delay bound D^*).



- o For each video source (*Jurassic Park, Star Trek*), use the empirical envelope E^* *to* compute N_{max} as a function of the delay bound D^* for the following values: D^* *=50, 100, 200, 300, 400, 500 ms*.

  Note: The value of N_{max} will increase with the delay bound.

- o For each type of video source, prepare a graph as shown above, which depicts that shows N_{max} as a function of D^*. The figure should include the values computed in Exercise 2.2 for the peak rate and mean rate allocation.

  Note: The value of N_{max} for the peak and mean rate allocations does not depend on D^*. The peak rate allocation will give a lower bound and the mean rate allocation will give an upper bound on the number of admissible flows.

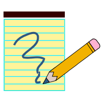
Exercise 2.3 Admission control for homogeneous traffic (Leaky bucket)

Repeat Exercise 2.2 with the leaky bucket traffic specifications from Exercise 1.3 as envelopes. Consider all three types of the leaky bucket:

- Leaky bucket;
- Peak-rate constrained leaky bucket;
- Dual-Leaky bucket.

- o Similar to Exercise 2.2, for each type of video source, prepare a graph that shows N_{max} as a function of D^* for each leaky bucket type. For comparison, include the results for the empirical envelope.

- o Modify the leaky bucket parameters to increase N_{max}. Provide a discussion of your findings.

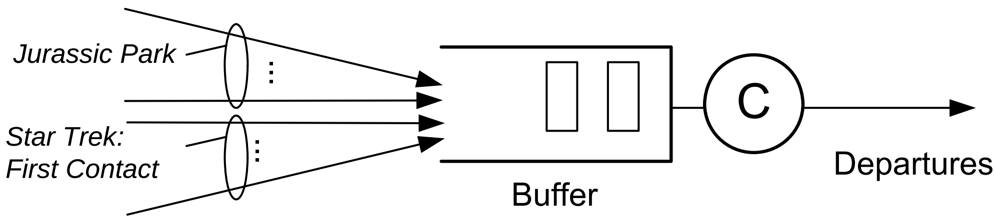


Lab Report:

Provide the plots and a discussion of the plots.

Part 3. Admissible Regions (Heterogeneous Traffic)

Consider a link (C=100 Mbps) with traffic from different types of flows with different delay requirements. Specifically, there are two types of video flows: *Jurassic Park* and *Star Trek: First Contact*. We refer to this arrival scenario as **heterogeneous traffic**.



The delay bounds of the flows are referred to as D_{JP}^* and D_{ST}^* , respectively, for the Jurassic Park and Star Trek video source, and are set as follows:

$$D_{JP}^* = 20\ ms$$
$$D_{ST}^* = 100\ ms$$

Consider two scheduling algorithms: FIFO and Earliest-Deadline-First (EDF). The objective of this lab is to explore how well these scheduling algorithms support the bounded delay service.

FIFO: The schedulability condition for the FIFO algorithm is as in Part 2, adjusted for the heterogeneous traffic scenario, is as follows:

$$\min\{D_{JP}^*, D_{ST}^*\} \geq \sup_{t\geq 0}\left\{\frac{1}{C}(N_{JP}\cdot E_{JP}(t) + N_{ST}\cdot E_{ST}(t) - Ct)\right\}$$

where

- N_{JP} and N_{ST} are the number of flows from each type;
- $E_{JP}(t)$ and $E_{ST}(t)$ are envelopes for each type of flow.

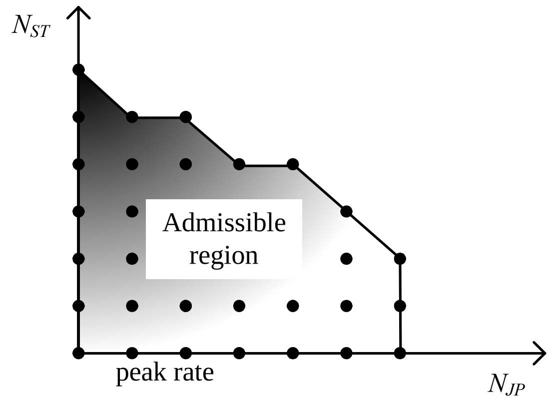
Since FIFO can only support one delay bound, the condition must specify that the smaller of the two delay bounds D_{JP}^* and D_{ST}^* is satisfied at all times.

EDF: The schedulability condition for the EDF scheduling algorithm (as derived in class), adapted to this scenario is:

$$\sup_{t\geq 0}\{N_{JP}\cdot E_{JP}(t - D_{JP}^*) + N_{ST}\cdot E_{ST}(t - D_{ST}^*) - Ct\} \leq 0$$

An admission control consists of verifying that the number of flows N_{JP} and N_{ST} satisfy the schedulability conditions. As long this is the case, no delay bound will be violated.

For given envelopes, given delay bounds and given link capacity, there will be certain pairs (N_{JP} , N_{ST}) that satisfy the schedulability condition of the (given) scheduling algorithms. The set of all pairs that satisfy the schedulability condition form an **admissible region**, as illustrated below. A set of flows consisting of N_{JP} *Jurassic Park* videos and N_{ST} *Star Trek* videos can be admitted, if the (N_{JP}, N_{ST}) lies inside the admissible region.



The objective of this part of the lab is to determine the admissible region of the FIFO and EDF scheduling algorithms for the given traffic scenario.

Exercise 3.1 Admissible region for mean and peak rate allocation

Consider a rate allocation scheme that allocates to each flow a capacity that is equal to the peak rate and the average rate of the video flows (as defined in Part 1). With all the parameters given above, determine the admissible regions for the peak rate allocation and the average rate allocation.

Hint: Suppose P_{JP} and P_{ST} denote the peak rates of the *Jurassic Park* and *Star Trek* video flows, respectively. Then (N_{JP}, N_{ST}) is in the admissible region if and only if

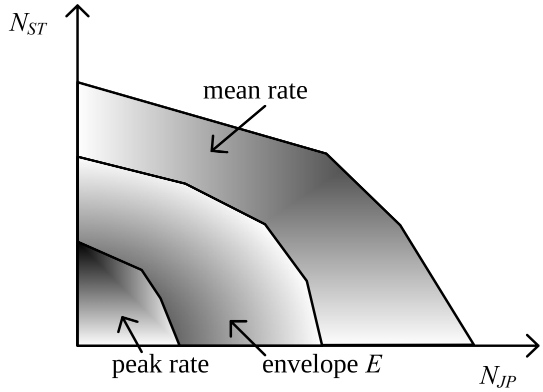
$$N_{JP} \cdot P_{JP} + N_{ST} \cdot P_{ST} \leq C.$$

The admissible region for the mean rate allocation is computed analogously.

- Provide a graph that depicts the admissible region of the peak rate and average rate allocation.

Exercise 3.2 Admissible region for FIFO and EDF (Peak rate-constrained leaky bucket)

With the network parameters, and schedulability conditions given above, compute the admissible region for the FIFO and EDF scheduling algorithms, and compare them to the results of Exercise 3.1.



- As envelope for this exercise, select the peak-rate constrained leaky bucket. Use the parameters that you determined in Exercise 2.3 for the *Jurassic Park* and the *Star Trek* video sources.

- Determine the admissible region for a link with a FIFO scheduling algorithm.

- Determine the admissible region for a link with an EDF scheduling algorithm.

- Prepare a graph that depicts the admissible region of both scheduling algorithms. Compare the results with the admissible regions from Exercise 2.3.

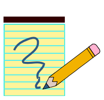
Exercise 2.4 (Optional, Up to 10% credit) More admissible regions

- Provide the admissible regions of FIFO and EDF for the following traffic specifications:

    - Leaky bucket
    - Dual-Leaky bucket
    - Empirical envelope

  The parameters of the traffic specifications are as obtained in7 Part 2 of this lab.

- Prepare graphs that depict and compare the admissible regions.



Lab Report:

Provide the plots and a discussion of the plots.

Feedback Form for Lab 4

- Complete this feedback form at the completion of the lab exercises and submit the form when submitting your lab report.

- The feedback is anonymous. **Do not put your name on this form** and keep it separate from your lab report.

- For each exercise, please record the following:

	Difficulty (-2,-1,0,1,2) -2 = too easy 0 = just fine 2 = too hard	**Interest Level (-2,-1,0,1,2)** -2 = low interest 0 = just fine 2 = high interest	**Time to complete (minutes)**
Part 1. Traffic specification of video traffic			
Part 2. Admission Control Tests (Homogeneous Traffic)			
Part 3. Admissible Regions (Heterogeneous Traffic)			

Please answer the following questions:

- What did you like about this lab?

- What did you dislike about this lab?

- Make a suggestion to improve the lab.