Delay Bounds for Scheduling Algorithms
The importance of a delay bound analysis for link schedulers is that it can be used to devise conditions that determine whether a scheduler can satisfy given delay requirements for a set of flows at a link scheduler with capacity $C$.

These conditions are called **schedulability conditions**.

A set of flows or traffic classes is said to be **schedulable** (with respect to a given scheduling algorithms) if all flows satisfy the schedulability conditions.
\[ N = \{1, 2, \ldots, N\} \] is set of flows

- Each flow \( j \) has a traffic envelope \( E_j \) (\( A_j \sim E_j \))

- **Goal:** Compute a condition so that the delay of any arrival does not exceed \( d \).
First-In-First-Out (FIFO)

- If we set
  \[ A(t) := \sum_{k=1}^{N} A_k(t) , \quad E(t) := \sum_{k=1}^{N} E_k(t) , \]
  we obtain a buffered link with a single arrival flow.

- The delay bound \( d^* \) is the smallest number \( d \) such that
  \[ \forall s \geq 0 : E(s - d) \leq Cs \]

- Substitute \( s - d \rightarrow s \):
  \[ d \geq \sup_{s>0} \left\{ \frac{E(s) - Cs}{C} \right\} \]
First-In-First-Out (FIFO)

Suppose

\[ E(t) := \sum_{k=1}^{N} (b_j + r_j t), \quad t \geq 0 \]

We get the FIFO delay bound:

\[ d^* = \frac{\sum_{k=1}^{N} b_j}{C} \]
$P$ priority classes ($1 =$ lowest, $P =$ highest)

Each class $p$ has a traffic envelope $E_p$ ($A_p \sim E_p$)

Ignore fact that packet transmission cannot be preempted.

**Goal:** Compute a condition so that the delay from priority $p$ does not exceed $d_p$. 
Consider a system with two priorities ($H$ = High, $L$ = Low)

\[ A_H \sim E_H \]
\[ A_L \sim E_L \]

Lower service curve for the low-priority traffic:

\[ S_L(t) = \left[ C \cdot t - E_H(t) \right]^+ \]

→ This is called the **residual service curve** or **leftover service curve**.
Static Priority (SP)

- Suppose

\[ E_H(t) = b_H + r_H t , \quad t \geq 0 \]
\[ E_L(t) = b_L + r_L t , \quad t \geq 0 \]
\[ C - r_H \geq 0 \]

- The delay bound \( d_L \) for low priority traffic is the smallest number \( d \) such that

\[ E_L \odot S_L(-d) \leq 0 \]

- This yields

\[ d_L = \frac{b_L + b_H}{C - r_H} \]
Back to multiple priorities. Suppose that for all $p$:

$$E_p(t) = b_p + r_p t, \quad t \geq 0$$

Mapping to a 2-priority system, we get for any class $p$

$$E_L(t) = b_p + r_p t, \quad t \geq 0$$

$$E_H(t) = \sum_{q=p+1}^{P} (b_q + r_q t), \quad t \geq 0$$

The SP delay bound $d_p$ ($d_p = d_L$) for class $p$ then is

$$d_p = \frac{\sum_{q=p}^{P} b_q}{C - \sum_{q=p+1}^{P} r_q}$$
Each class $j$ has a delay index $d_j$ (Assume: $i < j \Rightarrow d_i < d_j$)

Each class $j$ has a traffic envelope $E_j$ ($A_j \sim E_j$)

**Goal:** Compute a condition so that the delay for class $j$ does not exceed $d_j$

$\Rightarrow$ All class-$j$ traffic meets its deadline!
We skip the analysis of EDF, and only discuss results.

With EDF, the values for $d_1, d_2, \ldots$ are interdependent.

The delay bounds must be selected so that:

$$\sup_{s \geq 0} \left\{ \sum_k E_j (s - d_j) - Cs \right\} \leq 0$$
Suppose

\[ E_j(t) = b_j + r_j t \quad \forall j, \quad t \geq 0 \]

\[ \sum_j r_j \leq C \]

Then, the EDF delay bound for class \( j \) is

\[
d_j \geq \sum_{k=1}^{j} b_k - \sum_{k=1}^{j-1} r_k d_k \\
C - \sum_{k=1}^{j-1} r_k
\]