Bounding delay through a service element

- A source process $A(t)$ with envelope $E(t)$ is serviced by an element with lower service curve $S(t)$, giving departure process $D(t)$

<u>Theorem</u>: $D \geq A * \delta_{d_{max}}$

where $d_{max} = \inf\{d : E * \delta_d \leq S\}$

(i.e. the delay of process A through the service element is at most d_{max})

<u>Proof</u>:

$$
\begin{aligned}
D &\geq A * S \\
&\geq A * (E * \delta_{d_{max}}) \\
&= (A * E) * \delta_{d_{max}} \\
&\geq A * \delta_{d_{max}}
\end{aligned}
$$

$\square$

- Note: d_{max} is the least amount the envelope $E(t)$ needs to be shifted to the right s.t. it lies below the service curve $S(t)$

- Eg. $S(t) = ct^+$

$$E(t - d_{max}) \leq ct$$

$$\Rightarrow E(t) \leq c(t + d_{max})$$

$$\Rightarrow d_{max} = \sup_{t \geq 0} \left(\frac{E(t) - ct}{c} \right)$$

- To guarantee a delay of $\leq d$, need to reserve a minimum service rate of C_{min} where

$$\sup_{t \geq 0} \left(\frac{E(t)}{C_{min}} - t \right) = d$$

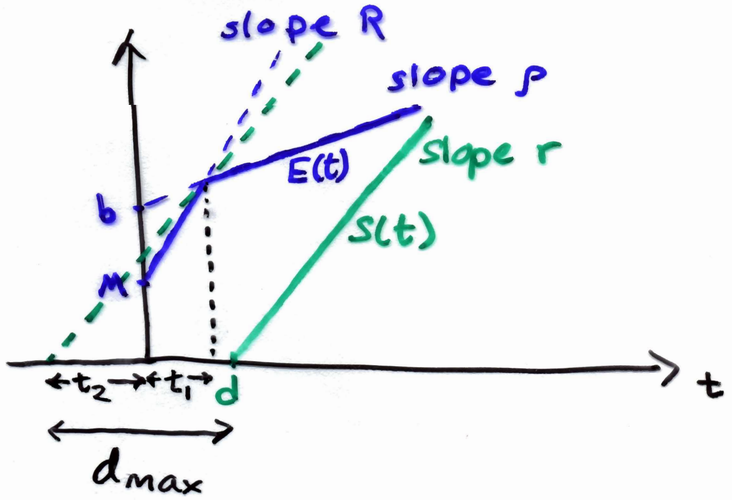
$$\Rightarrow \quad \frac{E(t)}{C_{min}} - t \leq d \quad \forall \; t \geq 0$$

$$\Rightarrow \quad C_{min} \geq \frac{E(t)}{d+t} \quad \forall \; t \geq 0$$

$$\Rightarrow \quad C_{min} = \sup_{t \geq 0} \left(\frac{E(t)}{d+t} \right)$$

- Eg. $E(t) = \min\left((M+Rt) I_{\{t \geq 0\}}, (b+\rho t) I_{\{t \geq 0\}}\right)$

$S(t) = r(t-d)^+$



$$M + Rt_1 = b + \rho t_1 \quad \Rightarrow \quad t_1 = \frac{b-M}{R-\rho}$$

$$r(t_1 + t_2) = M + Rt_1 \quad \Rightarrow \quad t_2 = \frac{1}{r}\left(M + t_1(R-r)\right)$$

$$= \frac{M}{r} + \frac{(b-M)(R-r)}{r(R-\rho)}$$

$$d_{max} = t_2 + d = \frac{M}{r} + \frac{(b-M)(R-r)}{r(R-\rho)} + d$$

- Eg. a connection with Tspec parameters (b, p, R, M) using a K-hop path where each hop uses store-&-forward & WFQ with reserved service rate r
  - let the max packet length over all connections at the jth hop be $L_{max,j}$, & the jth link's speed c_j
  - lower service curve at jth hop

$$r\left(t - \frac{L_{max,j}}{c_j} - \frac{M}{r}\right)^+$$

store-&-forward delay

  - end to end lower service curve

$$r\left(t - \sum_{j=0}^{K-1} \frac{L_{max,j}}{c_j} - \frac{KM}{r}\right)^+$$

  - total end to end delay

$$\frac{M}{r} + \frac{(b-M)(R-r)}{r(R-p)} + \sum_{j=0}^{K-1} \frac{L_{max,j}}{c_j} + \frac{KM}{r} + \sum_{j=0}^{K-1} T_{prop,j}$$

where $T_{prop,j}$ is the propagation delay at the jth link

- In IntServ, the Resource Reservation Protocol (RSVP) learns the parameters K, $L_{max,j}$, c_j, $T_{prop,j}$ of the path & determines the rate r needed for a connection with given Tspec parameters & required maximum end to end delay

- source initiates a connection by sending a PATH message containing its flowspec
- as the PATH message traverses the source-receiver path (assumed given),

$$\sum_{j=0}^{K-1} \left(\frac{L_{max,j}}{C_j} + T_{prop,j} \right)$$ & K are accumulated in the PATH message & connect^{n} state is established

- the receiver uses the info in the PATH message to calculate the required bandwidth to reserve, which is sent back in a RESV (RESerVe) message
- if at each hop sufficient bandwidth is available the reservation succeeds, otherwise a reservation failure message is generated

Superposition of source processes

- $A_i(t)$ has envelope $E_i(t)$, $1 \le i \le n$
- worst case deterministic envelope of the superposition of the n sources is $E(t) = \sum_{i=1}^{n} E_i(t)$

$$\left(\ge \sum_{i=1}^{n} \left(A_i(\tau + t) - A_i(\tau) \right) \right)$$

- eg. $E_i(t) = \min(Rt^+, (b+\rho t) I_{\{t \ge 0\}})$, $1 \le i \le n$
  - $\rightarrow E(t) = \min(nRt^+, n(b+\rho t) I_{\{t \ge 0\}})$
  - if $S_i(t) = r(t-d)^+$ gives a delay bound d_{max} for $E_i(t)$, then $S(t) = nr(t-d)^+$ gives delay bound d_{max} for $E(t)$

Multiplexer analysis with stationary & ergodic traffic

- Sources $i=1,2,\ldots,m$ bring amounts $B_k^{(i)}$ of data at times $t_k = kT$, $k \geq 0$ (discrete time model)
  - e.g. for a voice coder, T could be the coder's frame time (10–30 ms) & $B_k^{(i)}$ the number of bits in each frame (0 when source i is inactive)
- let X_k be the amount of data in the multiplexer buffer at time t_k
- let c be the multiplexer service rate

• Marginal buffering analysis

- multiplexer only stores packets arriving within a slot; packets not served within the slot are not carried over to the next slot
- relevant to the following scenarios:
  - stream sources with tight delivery deadlines
  - future all-optical packet networks with expensive optical buffering

- assume that for each i, the arrival process $\{B_k^{(i)}, k \geq 0\}$ is stationary & ergodic, & that the arrival processes of sources $i = 1, 2, \ldots, M$ are mutually independent
  - a random process $X_n, n \geq 0$ is stationary if $\forall\, k \geq 1$ & indices $n_1, n_2, \ldots, n_k$ & m, $(X_{n_1}, X_{n_2}, \ldots, X_{n_k})$ has the same distribution as $(X_{n_1+m}, X_{n_2+m}, \ldots, X_{n_k+m})$
  - X_n is ergodic if any invariant event (i.e. one which does not depend on any finite shift of the sequence) has probability 0 or 1
    $\Rightarrow$ time averages converge to expectations for an ergodic process (ergodic thm)
- e.g. Consider 2 biased coins with different probabilities p_1 & p_2 of heads. One coin is chosen at random with equal probability & tossed repeatedly. Let $X_n = 1$ if the nth toss gives heads & $X_n = 0$ otherwise. The sequence $X_1, X_2, \ldots$ is stationary but not ergodic

- for marginal buffering, stationary ergodic sources,
  - long run fraction of slots in which data is lost $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_{[\sum_{i=1}^{m} B_k^{(i)} > c]}$

$$= E\left(I_{[\sum_{i=1}^{m} B^{(i)} > c]}\right) \quad \text{w.p.1, by ergodic thm}$$

$$= Pr\left(\sum_{i=1}^{m} B^{(i)} > c\right)$$

  - long run fraction of data lost

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1}\left(\sum_{i=1}^{m} B_k^{(i)} - c\right)^+}{\sum_{k=0}^{n-1} \sum_{i=1}^{m} B_k^{(i)}} \quad \begin{array}{l} \leftarrow \text{data lost up to slot } n \\ \leftarrow \text{data arrived up to slot } n \end{array}$$

$$= \lim_{n\to\infty} \frac{\frac{1}{n}\sum_{k=0}^{n-1}\left(\sum_{i=1}^{m} B_k^{(i)} - c\right)^+}{\frac{1}{n}\sum_{k=0}^{n-1} \sum_{i=1}^{m} B_k^{(i)}}$$

$$= \frac{E\left(\sum_{i=1}^{m} B_0^{(i)} - c\right)^+}{E\left(\sum_{i=1}^{m} B_0^{(i)}\right)} \quad \text{w.p. 1}$$

  - long-run performance measures depend only on marginal distribution of data arrival process, not on the correlations across slots