Optimal Multiplexing on a Single Link: Delay and Buffer Requirements

IEEE Transactions on Information Theory
Volume.43 No.5 P.1518~P1535 Sep. 1997
Leonidas Georgiadis, Roch Guerin, Abhay Parekh

National Chiao-Tung University ,Communication Engineering
Name : Lin Shang-Yi
SID : 8913560
E_mail: ysl031@neta1.cm.nctu.edu.tw
Date : 89/11/29

Motivations(1/3)

- What it supports?
  - Multiple sessions with varying traffic characteristics.
  - Performance requirements in fast packet switched networks.

Motivations(2/3)

- Problems.
  - How to support a large number of sessions with different performance requirements.
  - Scheduling and buffer allocation policies.
    - Scheduling policy.
      - Determines the order in which queued packets are served.
    - Buffer allocation policy.
      - Determines the manner in which the buffer space is to shared among the sessions.
  - Consider network resources.
    - Link bandwidth .
    - buffer space(memory).

Motivations(3/3)

- Works.
  - Designing a scheduler.
    - Scheduling policy.
    - Buffer allocation policy.
      - Flexible Allocation(FL)
      - Semi-Flexible Allocation(SE)
      - Fixed Allocation(FI)
  - Tradeoff between efficiency and complexity.
  - Delay and buffer optimization.
  - "Soft" delay constraints.

Definitions(1/7)

- Multiplexer and traffic models.
  - Flow of each sessions partitioned into packets.
  - Packet can be arbitrarily small,but not longer than $L_{\max}$ bits.
  - Arriving packets are stored in the memory of the Multiplexer.
  - When the last bit of a packet has arrived then it become eligible for transmission.(*Store-and-forward* type)
  - Scheduler offer a *service policy* to decide which eligible packets to transmit in the output link.
  - Transmits this packet non-preemptively.(FIFO)

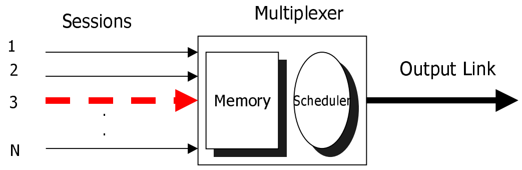

Fig.1

Definitions(2/7)

- Traffic constraints considered.
  - Packets average rate ρ_i .
  - Associated maximum burstiness σ_i .
  - Maximum packet size $L_{\max}$.
- Traffic generated by session i.
  - Number of bit.
  - $I_i(\tau, t+\tau) \le \sigma_i + \rho_i t \quad t \ge 0$ (1)
- Simple Assuming infinite input link speeds.
  - Amount of traffic from session i delivered to the scheduler.
  - $A_i(\tau, t+\tau) \le L_{\max} + \sigma_i + \rho_i t \quad t \ge 0$ (2)
- Stability assumption
  - $\sum_{i=1}^{N} \rho_i \le r$ (3)

Definitions(3/7)

- Some notation with implement policy π .
  - Number of session i bits served in the interval[τ ,t).
    - $S_i^{\pi}(\tau, t, \vec{A})$
  - Number of session i bits stored at time t.
    - $Q_i^{\pi}(t, \vec{A})$
  - Largest amount of bits from session i that can stored in the memory.
    - $C(\vec{\rho}, \vec{\sigma})$ the set of vectors of session traffic arrivals.
    - $\vec{A} \in C(\vec{\rho}, \vec{\sigma})$
    - $M_i^{\pi}(\vec{\rho}, \vec{\sigma})$
    - $M_i^{\pi}(\vec{\rho}, \vec{\sigma}) = \sup_{t \geq 0} \sup_{\vec{A} \in C(\vec{\rho}, \vec{\sigma})} Q_i^{\pi}(t, \vec{A})$

Definitions(4/7)

  - Delay of packet. The sum of time spent waiting in the memory.
    - Under traffic $\vec{A} \in C(\vec{\rho}, \vec{\sigma})$.
    - $D_i^{\pi}(\vec{\rho}, \vec{\sigma})$
- Notion of tracking policy $T(\pi)$.
  - The time at which packet p departs from a Multiplexer.
    - $f_p^{\pi}(t)$
  - Ordering Property.
    - Requires future arrivals do not modify the relative priorities of packets waiting to be transmitted.
    - Preemptive system f_p ,tracking system $\hat{f}_p$.
      - r is speed of input link sent to the Multiplexer
      - $\hat{f}_p - f_p < \frac{L_{\max}}{r}$. (4)

Definitions(5/7)

- Buffer Allocation Mechanisms.
  - Flexible allocation(FL)
    - $B_{FL}^{\pi} = \sup_{\vec{\rho}} \sup_{t \ge 0} \sup_{\vec{A} \in C(\vec{\rho},\vec{\sigma})} \sum_{i=1}^{N} Q_i^{\pi}(t,\vec{A})$ (5)
  - Semi-Flexible allocation(SE)
    - $B_{SE}^{\pi} = \sup_{\vec{\rho}} \sum_{i=1}^{N} \sup_{t \ge 0} \sup_{\vec{A} \in C(\vec{\rho},\vec{\sigma})} Q_i^{\pi}(t,\vec{A}) = \sup_{\vec{\rho}} \sum_{i=1}^{N} M_i^{\pi}(\vec{\rho},\vec{\sigma})$ (6)
  - Fixed Allocation(FI)
    - $B_{FI}^{\pi} = \sum_{i=1}^{N} \sup_{\vec{\rho}} \sup_{t \ge 0} \sup_{\vec{A} \in C(\vec{\rho},\vec{\sigma})} Q_i^{\pi}(t,\vec{A})$ (7)
  - Conclusion.
    - Buffer size using by above has $B_{FL}^{\pi} \le B_{SE}^{\pi} \le B_{FI}^{\pi}$.
    - Cost of implementation reduces from FL to SE to FI.

Definitions(6/7)

- Buffer requirements versus delay: Delay optimal policies.
  - Delay optimal policies.
    - Deadline
      - Delay vector (D1,D2,D3,…,DN)
      - Packet *p* arrive at time a_p .
      - Deadline : $d_p = a_p + D_i$.
    - Lateness
      - Finish time of packet p f_p .
      - Lateness : $l_p = f_{p-} - d_p$
    - Schedulable region.
      - Ω^{π}
      - Schedulable region admissible of policies C is $\bigcup_{\pi = C} \Omega^{\pi}$.
      - Region for PEDF and NPEDF.(Fig.2)
  - Delay optimal policies with Low buffer requirements.

Schedulable region

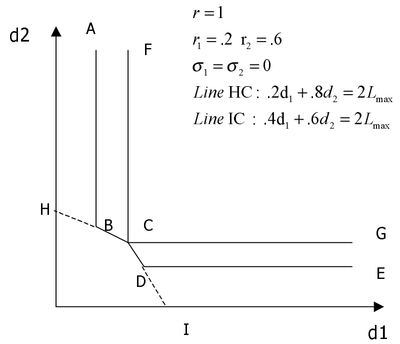

Fig.2

Definitions(7/7)

- Optimality criteria for Soft deadline.
  - Minimization of maximum lateness.
    - Above are assuming that the arriving traffic satisfies certain constraints.
    - Here keep the lateness of all packets as low as possible.
    - How do these policies performance change when deadline of packet are soft?
  - Lexicographic optimization.
    - The packet lateness vector is close to being *lexicographically* minimal.

Implementations

- Why it's useful?
  - Multiple session with varying traffic characteristic are discussed to see the performance of packet-switch networks.
  - Tradeoff between efficiency and complexity.
  - Use some buffer allocation policies and scheduling methods to improve the throughputs of the network.
  - From this analysis, we can make use of the network resources in the right way by using the above showed policies discusses.
  - Delays optimized will add flexibility results in significant advantages.

Issues

- Why use buffer allocation policies to improve the performance of the Multiplexer or said packet switch?
- Why use some scheduling policies in the output link of the Multiplexer?
- Delay taking into consideration and optimization.

Future Jobs

- Not only in Multiplexer here that should concern about the buffer issues, but also in other network devices.
- Now we can use these conclusions to think about the Delay problem in the network data delivering.
- Here discussed the non-preemptive service policy,which offered stored-and –forward type to receive the packets.Now we should consider some other issues,for examples VoIP ,which will change the priority of packet's order, and let voice packets to be the first priority.

Take Away

- Scheduling policies optimization in buffer and delay requirements.
- Buffer allocation policies tradeoff between efficiency and complexity.
- Delay optimal policies in PEDF, NPEDF, T(PEDF).
- "Soft" delay constraints provide for keeping the lateness of all packets as low as possible.

Related Concept(1/2)

- Optimality of PEDF for the class of preemptive policies.[19] For periodic arrivals and [11] for general arrival patterns.
- [16] , [17] the delay-optimality of *NPEDF* of non-preemptive policies for periodic and so-called sporadic arrivals.
- The schedulable regions for *NPEDF* and *PEDF* in [26] arrival streams with a minimum interpacket arrival time is independent of packet size.
- The merit of using schedulable regions to guarantee quality of service in networks was recognized in [18].
- The *NPEDF* policy in [12], [23], and [24] as a link scheduling policy provide per-session real-time guarantees in packet-switched networks.

Related Concept(2/2)

- Tracking policies was proposed and studied in Generalized Processor Sharing in [10] and [20].
- Theorems 1 and 2 appear in [20] but have been extended in this paper to include all tracking policies that obey a specific Ordering Property.
- Optimality of *PEDF* for the criterion of minimizing the maximum lateness of packets in [11].
- The relationship between *PEDF* and *NPEDF* in this context is new. The lexicographic optimality of *PEDF* is new as well.