

입력단에 버퍼가 있는 라우터를 위한 일정계획 방안

주운기*,이형섭**,이형호**

* 한국전자통신연구원 라우터기술연구부, 선문대학교 벤처 및 산업공학 전공

** 한국전자통신연구원 라우터기술연구부

A Scheduling Algorithm for Input-Queued Switches

Un Gi Joo*, H.S. Lee** and H.H. Lee**

* Router Technology Department, ETRI,

Department of Industrial Engineering, Sunmoon University

** Router Technology Department, ETRI

Abstract

This paper considers a scheduling algorithm for high-speed routers, where the router has an $N \times N$ port input-queued switch and the input queues are composed of N VOQ(Virtual Output Queue)s at each input port. The major concern of the paper is on the scheduling mechanism for the router. The paper discusses the preferred levels of the performance measures and then develops a non-linear mixed integer programming. Additionally, the paper suggests a heuristic scheduling algorithm for efficient and effective switching.

I. 서론

인터넷 사용자가 급증함에 따라 전송 링크의 용량 증설 및 고속 라우터의 필요성이 대두되고 있다. 대량의 데이터를 고속으로 처리하는 라우터를 위해서는, 입력 데이터별 정해진 출력포트로 원활하게 연결시켜 주기 위한 데이터별 처리 우선순위 결정 및 데이터전달(forwarding) 방안이 있어야 한다. 본 논문에서는 입력단에 VOQ(Virtual Output Queue)를 갖는 라우터에서의 단위 데이터(cell)별 입출력 포트 할당을 위한 일정계획 문제를 다룬다. 이러한 라우터 일정계획 수립

문제는 기존에 많은 연구가 진행되어 왔는데, 최적의 스위칭을 위한 일정계획 알고리즘을 위한 연구는 거의 없고 거의 대부분은 발견적(heuristic) 해법을 개발하는 것이다. 예를 들면, Mackeown and Anderson[4]은 입력단에 균등하고 독립적인 패킷이 입력된다고 할 때, Iterative round robin matching with slip(iSLIP) 방식의 발견적 알고리즘이 대기시간(queueing time)과 빈곤방지 면에서 좋다는 것을 시뮬레이션을 이용하여 보였다. 또한, Mekittikul and Mackeown[7]은 패킷의 대기시간을 weight로 하는 Maximal weight matching 방식의 발견적 방법이 트래픽 유형에 관계없이 100%의 수율(throughput)을 달성가능하고 빈곤(starvation)도 발생하지 않음을 보였다.

여러 유형의 발견적 알고리즘이 연구되고 있지만, 이들 알고리즘의 성능 비교 분석을 위한 객관적인 방법은 시뮬레이션을 포함한 수치적인 방법(numerical method) 외에는 아직 제시되고 있지 못하다. Bongiovanni et al.[1]이 수율최대화를 위한 알고리즘을 제시하였지만, 이것은 동적(dynamic)인 패킷 도착이 있는 상황에서는 사용할 수 없고, Mehmet-Ali et al.[6]의 분석 결과는 균등한 출력단 연결 요청이 아닌 경우에는 직접 이용할 수 없다.

본 논문에서는 라우터의 일정계획 알고리즘이 갖춰야 할 성질들이 어느 정도의 수준을 성능을 발휘하는 것이 최선인지를 결정한다. 그 다음으로는 각 성질 별 최선의 성능 수준을 가질 수 있는 알고리즘을 위한 수리계획(mathematical programming) 모형을 수립하여, 알고리즘의 성능 수준 평가를 위해 사용할 수 있는 분

석적인 방법(analytical method)을 제안한다. 마지막으로, 적절한 수준의 성능을 가질 수 있는 효율적인 일정계획 알고리즘을 하나 제시한다.

II. 최적 일정계획 수립

2.1 우선 서비스 순위 결정을 위한 고려사항

일반적으로 라우터를 위한 일정계획 알고리즘이 가져야 할 바람직한 성질로는 높은 수율(throughput), 빠른 처리, 빈곤(starvation) 방지, 구현(implementation)의 용이, 확장 및 조정의 용이성이 있어야 한다[2,5]. 그러나, 이들 척도 간에는 서로 상충(trade-off) 관계에 있는 것이 있으므로 이들 모두를 최선의 수준으로 달성할 수 있는 방안을 찾는 것은 매우 어렵다.

(1) 수율(throughput)

VOQ를 가지는 라우터에 대한 기존의 대부분의 연구는 입력 단에 도착하는 패킷에 의해 발생하는 제공부하(offered load)를 수율로 정의하였다. 그러나, Mekittikul and Mackeown[7]의 수율 정의에 의하면, 100%의 수율을 가질 수 있는 알고리즘이 그렇지 못한 알고리즘보다 나쁘지는 않겠지만, 이러한 알고리즘이 그렇지 못한 알고리즘보다 항상 더 나은 수율을 보장하는 것은 아니다. 더구나, 여기에서 사용한 100% 수율의 정의는 입력 패킷 간 상호 연관이 있는 집단도착(bulk arrival) 경우에는 이용할 수 없다. 따라서, 본 논문에서는 일반적으로 사용하는 수율의 척도인 다음과 같은 제공부하를 이용한다. 그러나, 이의 측정은 multicasting 또는 broadcasting 과 같은 서비스에 대해서도 적용할 수 있도록 출력단에서의 제공부하 ρ 를 수율로 측정한다.

$$\rho = \lambda \times E(S)$$

여기서, λ 는 단위시간당 출력단으로 연결이 설정된 패킷의 수를 나타내고, $E(S)$ 는 패킷 1개를 입출력 간 연결시키는 평균 시간(time slot)을 표시한다.

(2) 빠른 처리

패킷의 스위칭 시간 내에 다음에 스위칭을 할 패킷을 선정 완료하지 않으면, 스위칭 지연이 증가하고, 수율도 감소하므로, 스위칭이 처리되는 단위시간내에 알고리즘의 실행을 완료할 수 있어야 한다. 따라서, 이 알고리즘은 최소한 $O(N^a)$ 의 복잡도를 가지는 Polynomial complexity의 것이어야 한다. 여기서, 스위치의 단위시간 내에 최대 N^a 번의 연산을 수행한다고 할 때, 그 알고리즘의 복잡도 a 는 다음의 관계를 만족해야 한다.

$$a \leq \frac{\log(c/t)}{\log N}$$

여기서, N = 입력 포트 수

c = 단위 time slot으로, (한 패킷을 구성하는 총 비트 수)/(라인 처리속도 bps) second

t = 기본적 연산 1개를 처리 위한 시간

(3) 빈곤 방지

연결이 필요한 패킷이 존재하는데도, 특정 입력단의 패킷은 한번도 연결을 시켜주지 않는 형태의 알고리즘은 바람직하지 못하다. 이러한 현상은 스위치 전체적인 성능 향상만을 고려하는 경우 발생할 수 있는데, 스위치 전체적인 성능은 어느 입력단의 패킷에 대한 서비스를 해줄 것인가와는 관계없이 연결 설정을 기다리는 임의의 패킷에 대해 스위치가 쉬지않고 스위칭을 해주기만 한다면 수율은 최대가 될 것이기 때문이다. 이러한 사실은 다음과 같이 정리될 수 있다.

정리 1. 각 패킷별 연결설정(time slot) 시간이 동일하고, 연결 설정이 필요한 패킷이 존재하면 스위칭을 쉬지않고 수행한다면, 연결설정 우선순위에 관계없이 수율, 총 지연시간, 총 버퍼점유 수는 같다.

정리 1에 따르면, 수율, 지연시간, 버퍼점유 수 등의 스위치 전체의 성능만 고려한다면 특정 입력단의 패킷 서비스를 전혀 하지 않는 형태가 발생할 수 있다. 따라서, 빈곤 방지를 위해 본 논문에서는 평균 대기시간을 고려한다. 이외에 단일 칩(chip)내에 구현할 수 있는 정도로 알고리즘이 간단해야하고, 트래픽의 유형 및 서비스의 종류에 따라 우선 서비스 순위를 쉽게 조정할 수 있는 구조를 갖게하는 것이 좋다.

2.2 최적의 일정계획을 위한 수리계획 모형

이 장에서는 위에서 정의한 수율을 최대화하고 빈곤 방지를 위한 알고리즘을 위해 수리계획법 모형을 하나 제시한다. 이를 위해 다음의 기호를 먼저 정의한다.

$D_t = (D_{ijt})$: 시점 t 초의 입력단 i 에서 출력단 j 로 가기위해 대기중인 cell 수

$A_t = (A_{ijt})$: 시점 t 초에 입력단 i 에서 출력단 j 로 가기위해 도착하는 cell 수

$X_t = (X_{ijt})$: 입력단 i 에서 출력단 j 로의 시점 t 초에 연결 설정여부를 나타내는 0-1변수

여기서 서비스 시간을 나타내는 변수 t 는 time slot이 일정한 경우를 고려하므로, 1,2,3,... 등의 이산시간(discrete time)만 가지게되고, 단위 time slot 동안에 특정 입출력단 간 많아야 1개의 패킷만 연결을 설정가능하므로, 0또는 1의 값을 갖는 결정 변수 X_{ijt} 를 이용하여 수리계획 모형을 작성할 수 있다. 먼저, T

time slot 동안 출력단 j에서의 수율은 $\frac{\sum_{i=1}^N \sum_{t=1}^T X_{ijt}}{T}$

이 된다. 그리고, 입력단 i에서 출력단 j로 패킷이 단위 시간당 λ_{ij} 의 비율로 도착한다고 할 때, 입력단 i에서 time slot t에 실제로 출력되는 패킷흐름은 평균

$\sum_{j=1}^N \lambda_{ij} X_{ijt}$ 의 비율로 연결이 설정된다. 그러므로, M/D/1의 구조를 가정한다면, 입력단 i에서 출력단 j로의 패킷에 대한 평균 지연시간은

$$\frac{2 - \sum_{j=1}^N \lambda_{ij} X_{ijt}}{2(1 - \sum_{j=1}^N \lambda_{ij} X_{ijt})}$$

이 된다. 따라서, 수율은 최대화

하면서 평균지연은 최소화하기 위해서는 다음과 같은 수리계획 모형을 풀면된다.

$$\text{Max. } \frac{\sum_{i=1}^N \sum_{j=1}^N X_{ijt}}{T} - \frac{2 - \sum_{j=1}^N \lambda_{ij} X_{ijt}}{2(1 - \sum_{j=1}^N \lambda_{ij} X_{ijt})}$$

$$\text{s.t. } \sum_{i=1}^N X_{ijt} \leq 1$$

$$\sum_{j=1}^N X_{ijt} \leq 1$$

$$D_{ijt} = D_{ij,t-1} + A_{ijt} - X_{ijt}$$

$$X_{ijt} = 0 \text{ 또는 } 1, D_{ijt} \geq 0,$$

$$i,j=1,2,\dots, N, t=1,2,3,\dots, T$$

여기서, 세 번째 제약식은 현재 시점 t에서의 대기 패킷 수에 대한 균형식을 나타내는 것으로, 이는 각 입력단에 도착하는 패킷의 흐름을 나타내는 확률변수인 A_{ijt} 에 따라서 결정된다. 따라서, 이 모형을 풀기 위해서는 우선 각 입력단으로 들어오는 패킷 정보를 $t=1,2,\dots, T$ 동안 미리 발생시켜야 한다.

이 문제는 0-1 비선형 혼합정수계획법(0-1 nonlinear mixed integer programming) 문제이다. 이 문제를 풀기 위해서는 CPLX[3]와 같은 상용 패키지를 이용할 수 있다. 그러나, 일반적으로 해를 구하는데 걸리는 시간이 많이 걸리므로, 여기서 제안한 모형을 이용하여 실제로 패킷 일정계획을 수립하는데는 무리가 있다. 따라서, 다음 장에서는 빠른 수행을 위한 발견적(heuristic algorithm)을 하나 개발한다.

III. 발견적 일정계획 알고리즘

이 장에서 제시하는 알고리즘은 빠른 수행을 위해 반복적인 연산절차를 이용하고, 빈곤 방지를 위해서는 RRM(Round-Robin Matching) 방식을 취한다. 그리고,

수율 향상 및 빈곤 방지를 위해서는 OCF(Oldest Cell First) 우선순위 전략을 이용한다. 각 입출력단에서 서로 알고리즘의 자세한 절차는 다음의 기호에 대한 정의를 추가로 이용하여 기술한다.

d_{ij} = 입력단 i에서 출력단 j로 스위칭을 원하는 패킷 중 최대 대기시간

단계 1. (연결 요청) 각 입력단에서 연결을 기다리는 패킷 중 가장 오랫동안 연결이 설정되지 않았던 패킷에 대해, 그 패킷이 요구하는 출력단j를 향해 그 패킷이 기다린 시간 정보(d_{ij})와 함께 연결 요청 정보 보낸다.

단계 2. (연결 승인) 입력단으로부터 연결요청 메시지를 받은 각 출력단j는 연결 요청 메시지 중,

$$\text{Max}_{1 \leq i \leq N} \{d_{ij}\}$$

인 입력단을 향해 연결승인 정보를 보내고, 나머지 입력단에는 연결거부를 알리는 메시지를

보낸다. 여기서, $\text{Max}_{1 \leq i \leq N} \{d_{ij}\}$ 인 입력단이 두 개 이상인 경우에는 RRM 방식으로 할당을 위해 Pointer를 이용한다.

단계 3. (연결 선택) 연결승인 메시지를 받은 입력단에서는 연결승인을 받은 출력단으로 d_{ij} 의 대기시간을 가진 패킷을 연결시킨다.

여기서 제안하는 알고리즘은 'maximum size matching'과는 반복적이고 분산처리를 한다는 점에서 기본적인 차이가 있고, 'parallel iterative matching'과는 RRM 방식을 취하고 random selection이 없다는 점에서 차이가 있다. 또한 'iSLIP'과는 출력단에서의 연결승인 과정이 없고 패킷 우선순위를 결정하기 위해 OCF를 이용한다는 점에서 차이가 있다. 'random selection'을 위해서는 고속의 라우터를 구현하는데 복잡성을 증가하는 요인이 되므로 피하였고, OCF는 빈곤 방지 뿐 아니라 수율 측면에서도 좋은 기준이 될 수 있다는 Mekkittikul and Mekeown[7]의 결과를 이용하였다. 이를 이용해서 출력단에서는 입력패킷의 결정을 OCF를 기준으로하여 결정하므로 'iSLIP'보다 알고리즘이 단순화되어 빠른 수행을 할 수 있다.

여기서 제안한 알고리즘의 성능 평가를 위해서, 4×4 스위치 구조에 대한 다음과 같은 트래픽 D를 고려하자.

$$D = \begin{bmatrix} 1 & 2 & 4 & 2 \\ 3 & 2 & 2 & 4 \\ 1 & 3 & 1 & 1 \\ 5 & 2 & 4 & 1 \end{bmatrix}$$

현재 시점을 $t=5$ 라고 할 때, 각 입출력단을 위한 패킷들이 시점1에서부터 차례로 총 D_{ij} 개의 패킷이 도착

하여 현 시점의 각 입력단에서 대기 중인 패킷들이 $D=(D_{ij})$ 와 같다고 하고, 이후에는 추가로 도착되는 패킷들이 없다고 하자. 이러한 가정은 5 단위시간 동안 트래픽이 도착되다가 이후 일정시간 동안은 전혀 새로운 패킷이 도착되지 않는 불균형(non-uniform) 트래픽의 경우에 해당되므로 문제를 위한 심각한 가정은 아니다. 이러한 경우, 입력단 i 에서 출력단 j 로 연결하고자 하는 패킷의 수가 D_{ij} 라면, 주어진 트래픽 D 를 모두 스위칭하는 데에는 다음의 단위시간이 필요함을 쉽게 할 수 있다.

$$Max. \{ \lceil \frac{\sum_{i=1}^N \sum_{j=1}^N D_{ij}}{N} \rceil, \sum_{i=1}^N D_{ij}, \sum_{j=1}^N D_{ij} \}$$

여기서 $\lceil a \rceil$ 는 a 이상의 정수 중 가장 작은 정수를 의미한다. 따라서, 주어진 트래픽 D 에 대해서는 $Max \{ \lceil 38/4 \rceil, 10, 9, 11, 8, 9, 11, 6, 12 \} = 12$ 이므로, 최소한 12단위 스위칭 시간이 필요함을 알 수 있다.

이에 대한 처리를 위해 Mackeown[5]의 iSLIP 알고리즘을 이용하면, 14단위시간이 필요하다. 따라서, 이 문제에 대한 iSLIP 알고리즘의 수율은 $38/14 = 2.71$ 이 나온다. 반면에, 본 논문에서 제시하는 알고리즘에 의하면, 최소시간인 12단위 시간 내에 서비스를 완료하게 되어 수율 $38/12 = 3.17$ 을 가진다.

비록 이 논문의 알고리즘이 기존 알고리즘보다 나쁘지 않다는 것을 많은 예를 통해 보이지는 못하였지만 (이는 현재 진행중임), 본 논문의 알고리즘은 트래픽 유형에 관계없이 100%의 수율과 빈곤 방지를 할 수 있는 OCF 방식의 연결을 반복적으로 수행하는 방안을 택하고 있으므로 현재 수행중인 수치분석 결과가 좋을 것으로 예측된다.

IV. 결론

인터넷 등 통신 서비스 수요의 급증으로, 대량의 데이터를 고속으로 처리하는 라우터가 시급한 실정이다. 본 연구에서는 입력단에 VOQ(Virtual Output Queue)를 갖는 라우터에서의 단위 데이터(cell)별 입출력 포트 할당을 위한 일정계획 문제를 다룬다.

일반적으로 라우터를 위한 일정계획 알고리즘이 가져야 할 바람직한 성질로는 높은 수율, 빠른 처리, 빈곤 방지, 구현의 용이 등인데, 발생 가능한 모든 유형의 입력 데이터에 대해서 이와 같은 여러 가지의 성능 척도를 모두 만족스럽게 만족시킬 수 있는 발견적 해법을 개발하기는 쉽지가 않다. 본 논문에서는 라우터의 일정계획 알고리즘이 갖춰야할 성질들을 분석하여 각 성질 별 최선의 성능 수준을 가질 수 있는 알고리즘을 위한 수리계획(mathematical programming) 모형

을 수립하였다. 이는 라우터의 패킷 우선 서비스 순위 결정을 위해 직접 이용하기는 부적합하지만, 여러 알고리즘의 성능 수준 평가를 위해 사용할 수 있는 하나의 분석적인 방법(analytical method)으로 활용할 수 있다. 마지막으로, 적절한 수준의 성능을 발휘할 수 있는 효율적인 일정계획 알고리즘을 하나 제시하고, 수치 예를 통해 여기서 제안한 알고리즘이 최적의 성능에 가까운 것일 가능성이 있음을 보였다.

추후 연구로는 이 논문에서 제시한 수리계획법의 결과를 활용하는 알고리즘의 구현에 대한 연구가 필요한 상태이고, 여기서 제안하는 알고리즘의 성능 분석을 완성하는 것이다.

참고문헌

- [1] G. Bongiovanni, D. Coppersmith, and C.K. Wong, "An Optimum Time Slot Assignment Algorithm for an SS/TDMA System with Variable Number of Transponders", IEEE Transaction on Communications, Vol.COM-29, No.5, 1981, p.721-726.
- [2] M.R. Hashemi and A. Leon-Garcia, "The Single-Queue Switch : A Building Block for Switches with Programmable Scheduling", IEEE Journal on Selected Areas in Communications, Vol.15, No.5, 1997, p.785-794.
- [3] <http://www.cplex.com>, Mathematical Programming Software for Optimization, 2000.
- [4] N. Mackeown and T.E. Anderson, "A Quantitative Comparison of Iterative Scheduling Algorithms for Input-Queued Switches", Computers Networks and ISDN, Vol.30, No.24, 1998, p.2309-2326.
- [5] N. Mackeown, "The iSLIP Scheduling Algorithm for Input-Queued Switches", IEEE Transactions on Networking, Vol.7, No.2, 1999, p.188-201.
- [6] M.K. Mehmet-Ali, M. Youssefi and H.T. Nguyen, "The Performance Analysis and Implementation of an Input Access Scheme in a High-Speed Packet Switch", IEEE Transactions on Communications, Vol.42, No.12, 1994, p.3189-3199.
- [7] A. Mekittikul and N. Mckeown, "A Starvation-free Algorithm for Achieving 100% Throughput in an Input-Queued Switch", ICCN'96, 1996, p.226-231.