

High Speed Ethernet Switch 에서 COS기반의 Policy Daemon을 이용한 Dynamic Scheduler 에 대한 연구

최정일^o 황진옥 민성기
jeongil@korea.ac.kr

고려대학교 컴퓨터 과학기술 대학원

Study of Dynamic Scheduler that use Policy Daemon of COS base in High Speed Ethernet Switch

Jeong-il Choi^o Jin-ok Hwang Sung-gi Min
Dept. of Computer Science and Engineering . Korea Univ.

요 약

유선, 무선을 합하여 초고속 네트워크의 사용이 급격하게 증가함에 따라 고속의 네트워크 통신 서비스에 대한 요구가 증가 하였으며 차별화된 인터넷 서비스에 대한 요구가 증가함에 따라 IP 기반의 스위치가 등장하게 되었다. 이러한 스위치의 이슈는 IP 포워딩, 고속 스위치 패브릭 등의 문제를 포함하게 된다. 현재까지의 모든 패킷 스위치는 그 내부적으로 HOL 볼로킹에 의한 성능 저하의 영향을 최소화하기 위하여 여러 가지 스케줄링 알고리즘을 사용한다. 본 논문에서는 Gigabit 으로 동작하는 패킷 스위치 중에서 Strict Priority 알고리즘, Round Robin 알고리즘을 분석하고 각각의 스케줄러가 COS 기반과 어떤 관계가 있으며, 더불어 COS 기반에 따라, 스케줄러를 동적으로 변경하여 기존의 Static 하게 정책적으로 사용되던 Algorithm 과의 차별성을 두어 동적 스케줄러의 선택에 의한 Packet 처리 성능향상을 보이도록 한다.

1. 서 론

최근 네트워크에 연결한 사용자 수의 폭발적인 증가와 이 사용자들이 이용하는 멀티미디어 트래픽 의 양이 기하급수적으로 증가함에 따라 초고속 packet 을 처리해야 하는 Router 가 필요하게 되며, 이를 위해 고속 패킷 포워딩 기술 및 고속 패킷 스위칭을 위한 기술이 필요하게 되었다. 이러한 초고속 기술에 대해 보통의 Switch 의 경우 서로 다른 입력 단에 존재하는 여러 개의 셀들이 같은 목적지로 향하는 경우가 발생하게 되며 셀의 손실을 피하기 위하여 각각의 셀들을 저장하기 위한 큐잉을 필요로 하게 되었다.

이러한 큐잉중 입력큐잉은 HOL(Head of Line) 충돌 현상에 의한 Output 저하의 단점이 있게 되는데, 이러한 저하를 막고자 여러 가지 큐잉이 하드웨어적인 관점으로 고려되어 설계되어진다[1][2].

특히 Ethernet Switch 일 경우엔, 이러한 HOL 이 발생했을 경우 Ethernet packet 의 COS field 를 이용하여 packet 의 우선순위를 정하게 되는 경우가 보통이다. 여러 가지 H/W 적인 Scheduler 는 보통의 경우 Strict, Round Robin, Weighted Fair, 등의 Scheduler 가 있지만, 정책적인 결정은 보통 Router 나 Switch 가 운영자에 의해 세팅되어질 때 고정된다. 이렇게 고정된 값은 Dynamic 하게 변동

하는 요새의 Network 에 부합하기 힘들게 된다.

이러한 상태에 대하여 Dynamic 하게 Scheduler를 변동시킬 수 있는 Policy Daemon 을 두어 packet 을 분석하고 이것에 기반을 두어 Scheduler를 동적으로 세팅하게 하여 효율적인 네트워크 구성을 할 수 있음을 분석해 보았다.

2. 관련 연구

최근의 Ethernet 기반 스위치 및 라우터 들은 VLAN 기능을 기본적으로 내장하고 있다. 이 VLAN 기능 중에 일부가 COS 라 불리는 Class of Service 또는 Quality of Service 기능이라 할 수 있다 [3].

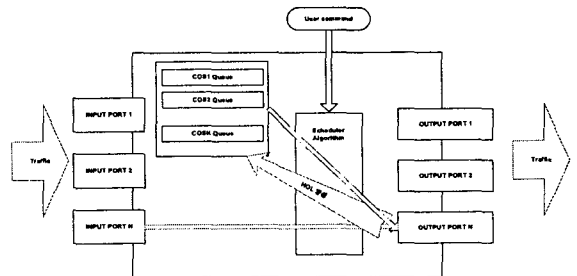


그림 1 HOL 발생 후 Scheduling 적용

이러한 기능들로 인하여 스위치들은 1에서 8까지의 CoS

queue 들을 내장하게 된다. 이러한 Queue 들은 각각의 traffic 에 대하여 스위치 장비들에 기본적으로 세팅된 scheduling algorithm 을 수행 하도록 되어있다. 기본적으로 많이 쓰이는 알고리즘은 Strict priority based queuing, Round Robin queuing, Weighted Fair Queuing Combination Queuing 이라 할 수 있다[4].

Strict priority based queuing 은 처리 되어야 하는 packet 은 그들의 priority 에 기반을 두어 고정되어 처리된다. Higher priority packet 의 경우 항상 lower priority packet 전에 처리되도록 되는 알고리즘이다. Round Robin queuing 은 각각의 priority 가 같은 weight로 설정되어 사용되는 알고리즘으로 모든 Queue 는 동일한 우선순위를 갖게 된다. Weighted Fair Queuing 은 각각의 Queue 에 대해 운용자의 설정에 따라 Program 되어지며 그 값에 따라 scheduling 되어진다. 물론 이때 실제 packet 이 Starvation 모드에 빠지지 않도록 운용자의 조정이 필요하게 된다. Combination Queuing 은 위의 3가지 Queuing 을 적절히 조정하여 쓸 수 있는 모드를 말한다. 특정 Queue 는 high priority 로 정하며 특정 Queue 는 Round Robin 모드로 정하여 사용하도록 하는 방법이다[5].

이와 같은 알고리즘은 [그림1]에서와 같이 HOL 이 발생하고 나면 각각의 Scheduling 에 의해 packet processing 이 진행된다.

3. 알고리즘 소개

기존의 방식은 위와 같은 알고리즘을 사용했을 경우 일반 운용자가 설정된 값으로 그 스위치 가 동작하도록 되어있다.

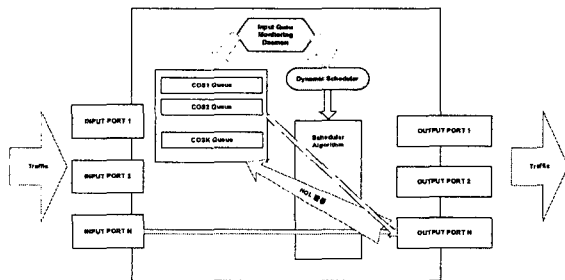


그림 2 HOL 발생 후 적용되는 Policy Daemon

서론 에서 언급한 것처럼 인터넷 packet 이 점점 고속화 되고 다양화 되어가는 이러한 구조에서 위와 같은 static 한 구조는 그 성능 면에 있어서 효율이 높지 않음을 알 수 있게 된다. 본 알고리즘은 실제로 들어오는 packet 의 양이 많

이 발생하여 HOL 가 발생할 경우 Scheduling 의 처리를 Dynamic 하게 변동 하도록 하는 Policy Daemon 을 구현한다[그림2].

Policy Daemon 알고리즘에서 Policy Daemon 은 실제 packet 과 HOL 의 여부를 검사하여 Dynamic 하게 스케줄링을 변경하도록 한다. Daemon 에는 Scheduler를 미리 세팅 하도록 하여 발생 가능한 정책을 위해 packet 검사를 위한 Threshold를 세팅하도록 한다. 모든 packet 에 적용될 수 있는 default scheduler 를 세팅하고, 이것을 Scheduler A 라 하고 해당 Scheduler 가 동작 할 수 있는 Threshold 역시 세팅 하도록 한다. 이와 동일한 방법으로 Scheduler B 와 Threshold 값을 세팅하도록 한다. Daemon 은 해당 스위치로 처리 되는 각각의 packet 을 분석토록하며 일정 시간 분석이 되면 그 분석치 와 함께 손실률을 계산하여 Scheduler 변경여부를 판단하게 된다.

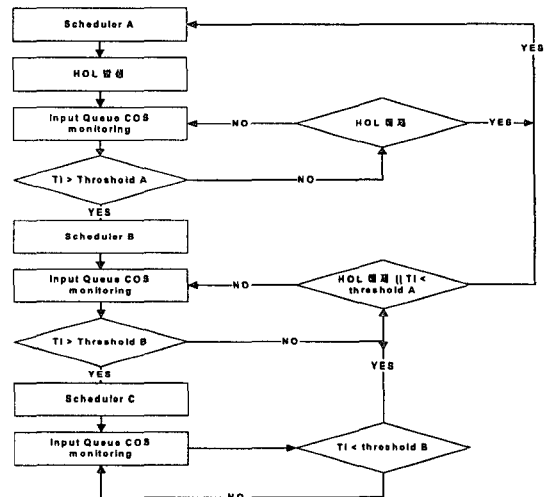


그림 3 Policy Daemon 알고리즘

처리율과 손실률은 입력된 packet 의 수와 설정된 COS Value, 그리고 Scheduler 와 의 관계에서 알아낼 수 있다. 이 Daemon 의 특정한 포트에서의 COS j 의 packet 처리율은 각각의 Queue 에 해당되는 Packet 의 수를 P_j 라 할 때 다음과 같이 계산 할 수 있다.

$$P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 = \sum P_i \quad (1)$$

$$P_j / \sum P_i = \quad (2)$$

더불어 입력이 uniform 하다고 가정하여 각각의 해당 큐에 대한 입력율을 l로 하여 계산한다면

$$l_0 + l_1 + l_2 + l_3 + l_4 + l_5 + l_6 + l_7 = \sum l_i \quad (3)$$

$$l_j / \sum l_i \quad (4)$$

위와 같은 식에 따라서 모든 Packet 의 처리는 l_j 의 변화

에 따라 처리율이 불변인 위의 (1) 식으로 인해 그 packet 의 손실이 발생 한다.

일반적인 HOL 이 발생하여 input 단의 처리율이 100%를 보장하지 못하고 단지 N%의 처리율을 보이면 (100-N)%의 packet 손실률을 보인다고 할 수 있게 된다.

이때 전체적으로 그 관계를 알아보면 입력 윌 $l_j/\sum l_i$ 에 대하여 처리율을 P_j/P_i 로 나타낼 수 있게 되며, 이러한 수치를 통하여 현재 들어오는 packet 의 입력량 과 해당 Cos queue 의 packet 처리율을 계산하여 손실률을 최소화 시키도록 Daemon Program 은 구성된다. 처음 세팅된 Scheduler 와 Threshold 값을 비교하여 해당 Threshold 값을 넘으면 다음 Scheduler 가 적용되도록 되며, 그 다음의 Scheduler 와 Threshold 값을 비교하여 다음 Scheduler를 적용하도록 한다.

4. 성능 평가

본 논문에서 제시한 알고리즘의 성능을 분석하기 위해서 기존의 고정된 Scheduler를 사용했을 경우의 packet 처리율과 Dynamic 하게 Scheduler를 변경 하여 사용 했을 경우의 packet 처리율을 비교해 보았다.

Packet 처리의 input 은 packet generator 인 Smatbit 6000B를 이용하였으며 output 은 packet 분석 툴인 AX4000를 이용 하였다. Main Switch chipset 은 Broadcom 사의 BCM5690 을 이용하였다.

성능 평가시의 제약 조건은 아래와 같다.

- 1) 모든 입력과 출력 단의 링크형성은 동일하다.
- 2) input 의 packet 은 COS value에 의존하여 정하였다.
- 3) Scheduling 와 HOL 을 적용하기 위하여 bandwidth 이상의 Packet 을 generation 한다.
- 4) static 한 결과가 worst 인 경우를 가정하여 두 개 Queue 의 Packet Flow를 만들었으며, 시험장비의 사용을 위해 packet injection 은 두 포트 에서만 generation 한다.

	input	weight	output	손실률
RoundRobin	90%		45%	45%
	10%		5%	5%
Weight	90%	1	40%	60%
	10%	9	10%	0%
Strict	90%		40%	60%
	10%		10%	0%

표 1 고정된 Scheduling Output

이와 같은 조건하의 시뮬레이션 결과는 static scheduler 가 worst 의 packet 의 경우를 감지하지 못한 경우 input packet 의 약 60%를 discard 시키는 결과를 가져온다.

이와 동일한 조건 하에서 packet monitoring을 실시하여 들어오는 Ethernet packet 의 COS value 가 worst 인 경우 COS value를 재조정하도록 하였을 경우 60% 보다 낮은 결과를 가져옴을 확인 할 수 있었다.

Threshold		weight	output	손실률	
55%	strict	90%		50%	40%
		10%		0%	10%
60%	weight	90%	9	49%	41%
		10%	1	1%	9%

표 2 Dynamic Scheduling Output

이때 Dynamic Scheduler Daemon 은 최적의 알고리즘 인 Strict 또는 Weighted Fair Queuing 을 선택하여 Static 하게 고정되어 있던 Scheduler 를 변경하여 그 효율을 높이기 하였다.

5. 결 론

위의 시험 결과 특정 COS 의 input packet 양이 특정 시간 이상 반복 될 경우 Scheduler를 Dynamic 하게 변경하도록 함으로써 Static 하게 Scheduler를 설정했을 때 보다 효율적인 Scheduler 가 선택 되어지고, 더불어 Dynamic scheduler 선택이 변화되는 network 망에 있어서 효율적임을 보일 수 있었다. Daemon 의 성능을 위해 향후 Daemon 의 packet 분석 방법, 적절한 Threshold 값을 찾는 방법 등의 연구가 계속 되어야 한다.

6. 참고 문헌.

- [1] J.G.Jim" Dai "The throughput of data switches with and without speed up"
- [2] T.Anderson, S.Owicki, J.Saxe, and C.Thacker"High Speed Switch Scheduling for Local Area Networks,"
- [3] IEEE 802.1q Standards fo local and metropolitan area networks: Virtual bridged local area networks. ACM Trans. Comput. Syst. pp.319-52,Nov.1993.
- [4] 변성혁, 이형호 "차세대 IP 스위치 및 라우터 기술"
- [5] "StrataXGS Theory of operations" BCM5690 Datasheet Broadcom
- [6] 정해영, 진재호, 이홍재, 최진규 "고속 패킷 스위치 스케줄링 알고리즘에 관한 연구" 한남 대학교.
- [7] 김충현, 손유익 "HOL 블로킹을 위한 그룹형 입력버퍼 ATM 스위치"