OBS망에서의 TCP 재전송을 고려한 Drop Policy의 성능 향상

김래영° 김현숙 김효진 송주석
연세대학교 컴퓨터과학과
{leon°, joelle, hyojin, jssong}@emerald.yonsei.ac.kr

Performance Improvement of the Drop Policy based on TCP Retransmission in Optical Burst Switched Networks

Laeyoung Kim° Hyunsook Kim Hyojin Kim Jooseok Song
Dept. of Computer Science, Yonsei University

요 약

OBS망에서 버스트의 충돌로 인한 버스트의 drop은 TCP의 성능에 중요한 영향을 끼치나, 기존의 drop policy에서는 이를 고려하지 않으며 TCP에 대한 연구로는 버스트의 assembling이 주를 이루고 있다. 본 논문에서는 OBS망에서 TCP의 재전송 문제를 drop policy와 연계하여 그 성능을 향상시키고자 한다. 본 논문에서 제안하는 drop policy는 버스트의 재전송 횟수가 더 작은 버스트를 drop시키는 TCP 기반 DP이다. TCP 기반 DP 모델과 일반적인 DP 모델의 성능을 ns-2를 이용한 시뮬레이션을 통해 평가하며, 이 때 시간의 변화에 따른 TCP throughput과 패킷의 drop rate을 비교 분석한다.

1. 서 론

해마다 급속도로 증가하는 인터넷 트래픽을 보다 빠르게 처리할 수 있는 진보된 인터넷 기술인 광 인터넷은 광 스위칭과 DWDM(Dense Wavelength Division Multiplexing) 기술로 인해 방대한 대역폭을 제공한다. 광 인터넷을 위한 스위칭 기술로는 Optical Circuit Switching(OCS), Optical Burst Switching(OBS) 그리고 Optical Packet Switching(OPS)이 있는데 그 중 OBS[1]는 OCS와 OPS의 장점을 취하고 단점을 보완한 기술이라고 할 수 있다.

OBS망에서 데이터전송의 기본단위는 버스트로, 버스트는 다수개의 데이터 패킷으로 구성되며 ingress node에서 assembling되어 OBS망으로 전송된다. 버스트를 수신한 egress node는 이를 다시 패킷으로 나누어 최종 목적지로 보낸다. Ingress node에서는 버스트에 앞서 컨트롤 패킷(또는 헤더)이 버스트의 전송경로를 설정하기 위해 보내지며, 버스트는 경로가 잘 설정되었다는 응답(acknowledgement)을 기다리지 않고 base offset time후에 전송된다[2]. 컨트롤 패킷을 받은 intermediate node에서는 버스트를 위하여 필요한 자원(대역폭)을 예약하게 된다. 버스트는 컨트롤 패킷에 의해 미리 예약된 경로를 통해 O/E/O의 변환을 전혀 거치지 않고 all-optical하게 전송된다.

OBS망에서 활발히 연구되는 사안으로 drop policy(DP)가 있다. OBS망에서 두개의 버스트가 동시에 동일한 링크상의 동일한 wavelength를 사용하려고 할 때 충돌(contention)이 발생한다. 충돌이 발생하면 충돌을 일으킨 버스트 중 하나만이 성공적으로 전송되고 나머지는 drop되는데, DP는 이 때 어떤 버스트를 drop시킬지를 결정하는데 필요한 정책이다. DP와 관련된 연구가 많이 있지만 TCP의 성능을 고려한 것은 존재하지 않는다.

TCP 성능과 관련해서는 주로 버스트의 assembling과 연계한 연구[3][4]가 있으나, 버스트의 drop 역시 TCP 성능에 영향을 끼치며 drop된 버스트가 적절히 재전송되지 않을 경우 TCP 성능에 저하를 가져올 수 있다. 이는 UDP와는 달리 TCP의 경우 drop되거나 에러가 발생한 패킷에 대해 재전송을 수행하며 이로 인해 congestion control algorithm[6]을 적용하는 특성에 기인한다. 이에 본 연구에서는 충돌이 발생한 경우 버스트의 재전송 횟수가 더 작은 버스트를 drop시키는 policy를 제안하고자 한다.

2장에서는 기존의 drop policy에 대하여 살펴보고 3장에서는 본 논문에서 제안하는 TCP 성능을 고려한 DP를 설명한다. 4장에서는 TCP 성능을 고려한 DP를 적용한 모델의 성능을 TCP throughput과 패킷의 drop rate을 토대로 평가한다. 마지막으로 5장에서는 결론을 맺는다.

2. OBS망에서의 DP(Drop Policy)

OBS망에서 하나의 버스트는 다수개의 패킷으로 이루어지므로, 일반적으로 그 크기가 패킷에 비하여 무척 크다고 볼 수 있다. 따라서 버스트가 하나만 손실되어도 망 전체의 loss rate에 중요한 영향을 미칠 수 있으므로, 버스트 자체의 survivability 뿐만 아니라 버스트간의 충돌 발생 시 어떤 버스트를 drop시킬 것인지를 결정하는 DP가 중요한 이슈가 될 수 있다.

기존의 연구에서는 버스트의 충돌을 해결하고자 다양한 메커니즘들이 연구되어 왔으며, 이와 더불어 적절한 DP가 사용되어졌다. 충돌 해결 메커니즘에 대한 내용은 이 논문의 범위를 벗어나므로 여기에서는 생략하기로 한다. 기존의 연구에서 보이는 DP들은 명백하게 분류되어 있거나, 정의되어 있지는 않지만, 크게 일반적인 DP와 priority 기반 DP로 나누어 볼 수 있다.

2.1 일반적인 DP

이 방식은 충돌이 발생했을 때, 일반적으로 시간상 먼저 채널을 점유하게 되는 버스트(original burst)를 살리고, 나중에 도착하여 충돌을 발생시킨 버스트(contending burst)를 drop시

키는 정책을 말한다. 이는 일반적인 drop-tail queue와 유사한 방식이라고 볼 수 있다.

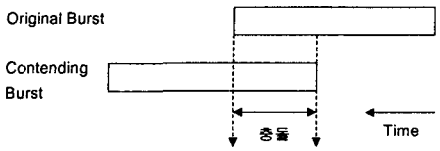

그림 1 충돌

2.2 Priority 기반 DP

이 방식은 충돌을 일으킨 버스트들의 priority를 고려하여 drop시킬 버스트를 결정하는 방식으로 priority는 ingress node에서 부여할 수도 있고, intermediate node에서 네트워크의 상태에 따라 이를 조정하는 메커니즘도 있다.

[2]에서 제안한 offset-time 기반 QoS 방안에서는 ingress node에서 QoS class에 따라 base offset time과 구분되는 extra offset time을 추가로 설정한다. 이 때 버스트의 priority가 클수록 extra offset time에 더 큰 값을 지정함으로써 class isolation을 구현하다. 이 방안에서는 extra offset time이 drop될 버스트를 결정짓는 priority로 작용한다.

[5]에서 제안하고 있는 LDR(Limited Deflection Routing)은 충돌을 해결하고자 deflection routing을 보완한 메커니즘이며, priority 기반 DP를 사용한다. 여기서는 버스트의 priority를 결정짓는 요인으로 충돌이 발생한 intermediate node로부터 egress node까지의 남은 홉 수, 버스트가 앞으로 지나가게 될 path의 평균 blocking probability 그리고 버스트의 길이 등을 언급하고 있다. 이와 같은 요소들을 priority의 기준으로 정의한 이유는 궁극적으로 네트워크의 전체 버스트 loss rate을 줄이기 위한 노력의 하나라고 할 수 있다.

3. TCP 성능을 고려한 Drop Policy

OBS망에서 TCP와 관련된 기존의 연구에서는 주로 ingress node에서 버스트가 assembling될 때 소요되는 delay가 TCP의 성능에 영향을 줌을 보이고 있다[3][4]. 그런데 버스트의 assembling에 걸리는 delay 뿐만 아니라 버스트의 drop 역시 TCP의 성능에 영향을 끼치는 요소이다. 만약 drop된 버스트에 대한 재전송이 적절히 이루어지지 않을 경우, 패킷을 보낸 소스단의 TCP는 이로 인해 congestion control algorithm[6]을 수행할 것이며 이는 TCP의 성능에 저하를 가져올 수 있다. 이에 본 논문에서는 충돌이 발생하여 drop시킬 버스트를 선택해야 할 때 버스트의 재전송 횟수가 priority로 작용하는 DP를 제안하고자 한다.

본 논문에서 제안하는 TCP 기반 DP를 위해서는 버스트의 컨트롤 패킷에 "Retransmission Count"라는 필드가 추가되어야 한다. Ingress node에서 버스트를 assembling한 후 이를 OBS망으로 맨 처음 전송하는 경우, 이 값은 0으로 설정되어야 한다.

컨트롤 패킷이 버스트를 위한 전송경로를 예약하기 위해 OBS망을 지나다가 어떤 intermediate node에서 충돌이 발생한 경우, 그 노드는 drop되는 버스트에 대해서 ingress node로 NAK 메시지를 보냄으로써 경로예약이 실패했음을 알리게 된다. 이러한 NAK 메시지에도 "Retransmission Count" 필드가 포함되어야 한다.

본 논문에서는 일반적인 버스트의 priority는 존재하지 않는다고 가정하며, 그림 2에서 보여주는 TCP 기반 DP의 동작을 자세히 살펴보면 다음과 같다.

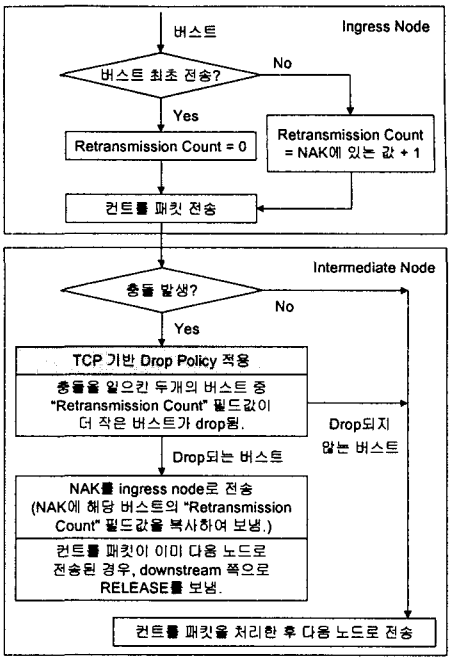

그림 2 TCP 기반 DP의 동작

1. Ingress node가 처음으로 전송되는 버스트의 컨트롤 패킷에 있는 "Retransmission Count" 필드값을 0으로 설정하여 컨트롤 패킷을 OBS망으로 전송한다.
2. Intermediate node에서 충돌이 발생한 경우, TCP 기반 DP를 적용하여 drop시킬 버스트를 선택한다. 즉, 충돌을 일으킨 두개의 버스트에 대해 "Retransmission Count" 필드값을 비교하여, 그 값이 더 작은 버스트를 drop시킨다. 만약 "Retransmission Count" 필드값이 서로 같다면 일반적인 DP처럼 original burst를 계속 전송하고, contending burst를 drop시킨다.
3. Drop되는 버스트에 대한 처리를 다음과 같이 수행한다.
   1) Drop되는 버스트에 대해서는 충돌이 발생한 intermediate node에서 ingress node로 경로예약이 실패했음을 알리는 NAK 메시지를 보낸다. 이 때 NAK 메시지의 "Retransmission Count" 필드값은 해당 버스트의 컨트롤 패킷에 있는 값을 복사한 것이다.
   2) Drop되는 버스트의 컨트롤 패킷이 이미 충돌이 발생한 노드를 지나서 다음 노드로 전송된 경우, downstream 쪽으로 RELEASE 메시지를 보내어 이미 예약된 자원을 해제한다.
4. NAK 메시지를 받은 ingress node는 재전송해야하는 버스트의 컨트롤 패킷의 "Retransmission Count" 필드값을 NAK 메시지에 있는 "Retransmission Count" 필드값 + 1로 설정하여 컨트롤 패킷을 다시 전송한다.
5. Intermediate node에서 충돌이 발생하지 않았거나, 충돌로 인해 DP를 적용하여 계속 전송할 버스트를 선택한 후에는 해당 버스트의 컨트롤 패킷을 처리하여 자원을 예약한 다음 이를 다음 노드로 전송한다.

물론 이 동작에서 ingress node는 컨트롤 패킷을 전송한 후 base offset time이 지나면 해당 버스트를 전송하게 된다.

4. 성능 평가

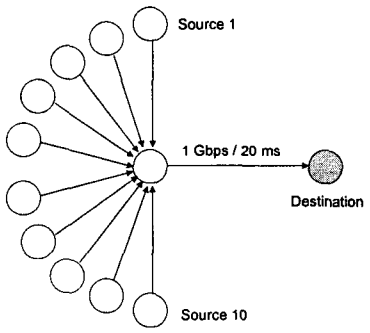

그림 3 시뮬레이션 토폴로지

본 논문에서 제안하는 TCP 기반 DP를 적용한 모델과 기존의 일반적인 DP를 적용한 모델의 성능을 비교, 평가하기 위해 ns-2[7]를 사용하여 시뮬레이션하였다. 그림 3은 시뮬레이션 토폴로지를 보여준다[8]. 이는 TCP의 성능을 평가하기 위해 OBS망을 단순화한 토폴로지로 10개의 source가 TCP 및 UDP 트래픽을 destination으로 전송하며, TCP의 경우 각 source와 destination 쌍마다 session을 2개씩 생성하였다. TCP의 경우 ftp 트래픽이 전송되며, UDP의 경우 exponential하게 분포된 트래픽이 전송된다. 버스트 하나의 크기는 1 Mbyte로 동일하다.

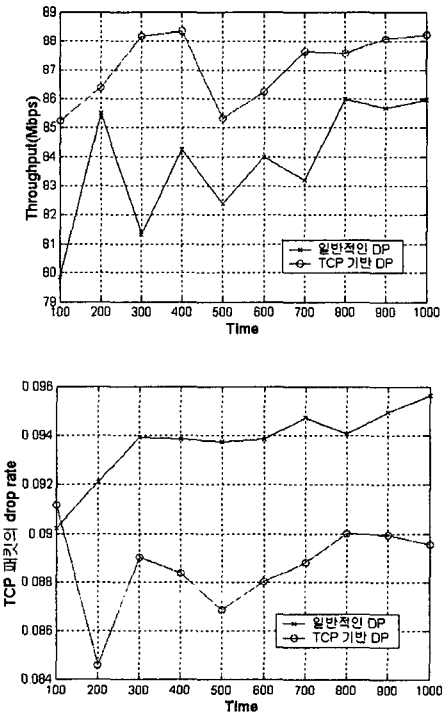

그림 4는 기존의 일반적인 DP와 제안하는 TCP 기반 DP를 적용한 모델의 시간 변화에 따른 TCP의 throughput을 나타낸 그래프이다. 시뮬레이션 타임(sec)이 증가함에 따라 destination이 초당 받은 TCP 패킷의 양을 Mbps로 나타낸 이 그래프는 일반적인 DP 모델이 평균 83.82 Mbps를 받는 반면, TCP 기반 DP 모델은 평균 87.13 Mbps로, 제안한 방식이 평균적으로 약 3.3 Mbps의 성능향상을 보이며, 최고 6.84 Mbps의 차이를 나타낸다.

그림 5는 두 모델에서 나타난 TCP 패킷의 drop rate을 보여주고 있다. 평균 8.8×10^{-2}의 drop rate을 보이는 TCP 기반 DP 모델이 9.3×10^{-2}의 drop rate을 보이는 일반적인 DP 모델에 비해 약 5% 더 낮은 drop rate을 나타내고 있다.

마지막으로 시뮬레이션 결과 두 모델의 패킷당 평균 end-to-end delay는 일반적인 DP의 경우 0.302 sec, TCP 기반 DP의 경우 0.319 sec를 보임으로써, TCP 기반 DP 모델이 약 0.0168 sec 정도 더 지연됨을 보이나, 이는 고속의 광 인터넷망의 특성상 수용할 만한 차이라 할 수 있다.

5. 결론 및 향후과제

본 논문에서는 TCP의 특성을 고려하여 OBS망에서 버스트의 재전송 횟수에 기반을 둔 DP를 제안하고 있으며, 이는 OBS망에서 TCP가 보다 효율적으로 동작할 수 있도록 한다. 버스트의 충돌 발생 시 재전송 횟수가 더 큰 버스트가 더 높은 priority를 갖는 것으로 간주하는 TCP 기반 DP는 일반적인 DP와 비교하여 TCP의 throughput을 높일 뿐 아니라, TCP 패킷의 drop rate 측면에서 또한 보다 좋은 성능 평가 결과를 보인다. 그리고 평균 end-to-end delay는 OBS망이 수용할만한 작은 차이를 보인다.

향후 연구 과제로, 본 논문에서 제안한 TCP 기반 DP를 수많은 소스로부터 영향을 받게 되는 버스트의 assembling 메커니즘과 연계하여 OBS망에서의 TCP 성능을 더욱 높일 수 있도록 개선하고자 한다.

6. 참고 문헌

[1] C. Qiao, M. Yoo, "Optical burst switching(OBS)-A new paradigm for an optical internet," J. High Speed Network, vol.8, pp.69-84, 1999.

[2] M. Yoo, C. Qiao, S. Dixit, "Optical Burst Switching for Service Differentiation in the Next-Generation Optical Internet," IEEE Commun. Magazine, pp.98-104, Feb. 2001.

[3] S. Gowda, R.K. Shenai, K.M. Sivalingam, H.C. Cankaya, "Performance evaluation of TCP over optical burst-switched(obs) wdm networks," Proc. ICC '03, vol.2, pp.1433-1437, May 2003.

[4] A. Detti, M. Listanti, "Impact of segments aggregation on TCP Reno flows in optical burst switching networks", Proc. INFOCOM 2002, vol.3, pp.1803-1812, June 2002.

[5] H.S. Kim, S.K. Lee, J.S. Song, "Optical Burst Switching with Limited Deflection Routing Rules", IEICE Trans. Commun. Vol.E86-B, No.5, pp.1550-1554, May 2003.

[6] M. Allman, V. Paxson, W. Stevens, "TCP Congestion Control," RFC 2581, April 1999.

[7] The Network Simulator - ns-2
http://www.isi.edu/nsnam/ns/

[8] S. Y. Wang, "Using TCP Congestion Control to Improve the Performances of Optical Burst Switched Networks," Proc. ICC'03, vol.2, pp.1438-1442, 2003.