



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0009064

(43) 공개일자 2022년01월24일

(51) 국제특허분류(Int. Cl.)

H04L 47/31 (2022.01) H04L 47/00 (2022.01)

H04L 47/6275 (2022.01)

(52) CPC특허분류

H04L 47/2458 (2013.01)

H04L 47/28 (2013.01)

(21) 출원번호 10-2020-0087306

(22) 출원일자 2020년07월15일

심사청구일자 2020년07월15일

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

백상현

서울특별시 서초구 고무래로 35, 109동 1004호

최진국

대한민국, 10351, 경기도 고양시 일산서구 고양대로 633, 107동 2402호

임용재

대한민국, 12582, 경기도 양평군 양서면 대심길 52-3

(74) 대리인

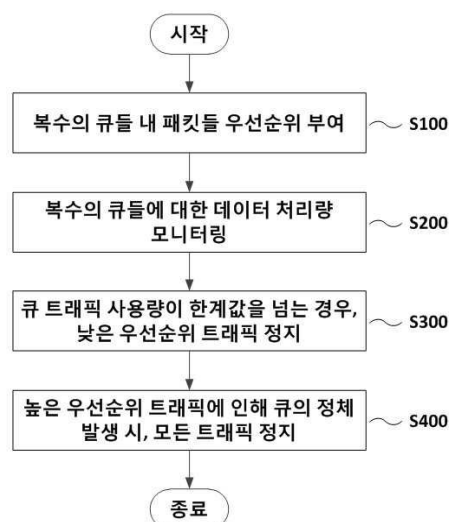
김홍석

전체 청구항 수 : 총 6 항

(54) 발명의 명칭 이더넷 기반 네트워크 내 트래픽 제어 방법

(57) 요약

이더넷 기반 네트워크 내 트래픽 제어 방법이 개시된다. 이더넷 기반 네트워크 내 트래픽 제어 방법은 복수의 큐들에 대한 데이터 처리량을 모니터링하고, 각 큐 내 패킷들에서 저지연을 요구하는 데이터 흐름의 패킷에 높은 우선순위를 부여하며, 복수의 큐 중 적어도 하나가 혼잡 시, 해당 큐 내 낮은 우선순위의 패킷의 흐름을 일시적으로 정지하고 높은 우선순위의 패킷을 우선 전송하는 단계를 포함하되, 정지된 패킷은 {C, T, R}의 집합으로 표기되며, C는 혼잡한 큐이며, T는 낮은 우선순위의 패킷이 정지하는 시간이며, R는 정지할 트래픽의 우선순위 최고 랭크 값이며, T 및 R 각각이 능동적으로 조절 가능하다.

대표도 - 도1

(52) CPC특허분류

H04L 47/6215 (2013.01)

H04L 47/6275 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711102841
과제번호	2017-0-00195-004
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	방송통신산업기술개발(R&D)
연구과제명	멀티 서비스를 지원하는 프로그래머블 스위치 제어 기술 개발
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

각각이 CoS(Class of Service) 타입에 대응하는 복수의 큐들이 구비된 네트워크 내의 스위칭 장치에서 수행되는 트래픽 제어 방법에 있어서,

상기 복수의 큐들 각각에 대한 데이터 처리량을 모니터링하는 단계; 및

상기 복수의 큐들 중 어느 하나의 큐에 혼잡이 발생한 경우, 적어도 하나의 패킷의 전송 정지를 요청하는 정지 프레임을 적어도 하나의 상위 스위칭 장치에 전송하는 단계를 포함하고,

상기 정지 프레임은 혼잡이 발생한 상기 어느 하나의 큐에 대응하는 CoS 타입을 식별하기 위한 정보, 패킷의 전송을 정지할 정지 기간에 대한 정보, 및 정지할 트래픽의 우선순위 최고 랭크 값을 포함하는,

트래픽 제어 방법.

청구항 2

제1항에 있어서,

상기 모니터링하는 단계는, 상기 어느 하나의 큐의 사용량이 제1 한계값을 초과하는 경우 상기 어느 하나의 큐에 혼잡이 발생한 것으로 판단하는,

트래픽 제어 방법.

청구항 3

제1항에 있어서,

상기 정지 프레임을 상기 적어도 하나의 상위 스위칭 장치에 전송하는 단계 이후에,

상기 적어도 하나의 상위 스위칭 장치로부터 상기 우선순위 최고 랭크 값 보다 높은 우선순위의 트래픽을 수신하는 단계를 더 포함하는,

트래픽 제어 방법.

청구항 4

제3항에 있어서,

상기 어느 하나의 큐에서 상기 높은 우선순위의 트래픽으로 인하여 정체가 발생한 경우, 상기 적어도 하나의 상위 스위칭 장치에 모든 트래픽의 전송 정지를 요청하는 제2 정지 프레임을 전송하는 단계를 더 포함하는,

트래픽 제어 방법.

청구항 5

제1항에 있어서,

정지할 트래픽의 우선순위 최고 랭크 값(R)이 m (m 은 음이 아닌 자연수)인 경우, m 보다 큰 값의 우선순위를 갖는 패킷에 대한 전송의 정지를 의미하며,

상기 R이 0인 경우, 상기 어느 하나의 큐로의 모든 패킷 전송의 전송의 정지를 의미하며,
상기 정지 기간에 대한 정보(T)가 0인 경우, 모든 패킷에 대한 전송 정지의 해제를 의미하는,
트래픽 제어 방법.

청구항 6

제1항에 있어서,
상기 복수의 큐들 각각은 PIFO(Push-in First-out) 방식을 사용하는,
트래픽 제어 방법.

발명의 설명

기술 분야

[0001] 본 발명은 네트워크 내 트래픽 제어 방법에 관한 것으로, 보다 상세하게는 이더넷 기반 네트워크 내 트래픽 제어 방법에 관한 것이다.

배경 기술

[0002] 최근 5G 네트워크를 기반으로 하는 네트워크 애플리케이션들에서 요구하는 고성능, 저지연 등에 대한 서비스 수준 협약(service level agreement, SLA)을 보장하기 위하여 데이터 센터 인프라에서는 무손실 이더넷 기반의 네트워크 프로토콜을 사용하여 서버 간의 데이터를 송/수신한다.

[0003] 우선순위 기반 흐름 제어(priority-based flow control, PFC)는 무손실 이더넷을 지원하기 위한 대표적인 흐름 제어 기법이다. 그러나, 기존 PFC 기법에서는 지원 가능한 CoS(Class of Service) 타입의 종류가 제한적이다. 트래픽에 대하여 흐름 제어를 진행하는데 이는 데이터 센터에서 실시간으로 발생하는 대규모의 플로우를 매핑하고 처리하기에는 제한적이다. 데이터 센터의 각 트래픽 플로우를 분석한 결과, 저지연을 요구하는 퀴리 트래픽이 파일 전송 트래픽으로 인하여 정체가 발생한 것으로 확인되었다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 대한민국 등록특허 제10-0674727호

발명의 내용

해결하려는 과제

[0005] 본원 발명이 해결하고자 하는 과제는 큐 혼잡 시 보다 효율적으로 트래픽을 제어하는 이더넷 기반 네트워크 내 트래픽 제어 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 해결하고자 하는 과제를 달성하기 위하여 본 발명의 실시 예들에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법은, 복수의 큐들에 대한 데이터 처리량을 모니터링하는 단계, 각 큐 내 패킷들에서 저지연을 요구하는 데이터 흐름에 대하여 높은 우선순위를 부여하는 단계, 및 상기 복수의 큐 중 적어도 하나가 혼잡 시, 낮은 우선순위의 패킷의 흐름을 일시적으로 정지하고 높은 우선순위의 패킷을 우선 전송하는 단계를 포함하되, 상기 정지된 패킷은 {C, T, R}의 집합으로 표기되며, C는 상기 혼잡한 큐이며, T는 상기 낮은 우선순위의 패킷이 정지하는 시간이며, R는 상기 정지할 트래픽의 우선순위 최고 랭크 값이며, 상기 T 및 R 각각이 능동적으로 조절 가능하다.

발명의 효과

- [0007] 본 발명에 의한 경우, 큐 혼잡을 탐지 시, 낮은 우선순위 패킷을 우선적으로 정지하여, 높은 우선순위 패킷은 낮은 우선순위 패킷에 비해 상대적으로 낮은 큐 지연 및 높은 대역폭을 가진다.
- [0008] 또한, 낮은 우선순위 트래픽에 대한 정지가 발생 시, 새로 생성된 높은 우선순위 트래픽에 대해서는 정상적으로 출력한다. 그리고 높은 우선순위 트래픽에 인하여 큐 혼잡이 발생 시, 전체 트래픽을 정지하는데, 높은 우선순위 트래픽의 대역폭이 해당 큐의 출력 속도보다 낮을 경우, 높은 우선순위 트래픽의 입력 대역폭과 같은 출력을 보장한다.
- [0009] 같은 CoS 값을 가지는 높은 우선순위 트래픽과 낮은 우선순위 트래픽에 대하여, 높은 우선순위 트래픽이 상대적으로 낮은 큐 지연 및 높은 대역폭을 가질 수 있다. 최종적으로 높은 우선순위 트래픽이 더 낮은 FCT(Flow Completion Time)를 가질 수 있다.
- [0010] 또한, 다중 지정량에 대한 설정을 통하여, 능동적으로 큐 혼잡제어가 가능하다.

도면의 간단한 설명

- [0011] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.
- 도 1은 본 발명의 일 실시예에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 위한 순서도이다.
- 도 2는 본 발명의 일 실시예에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 블록도이다.
- 도 3은 본 발명의 일 실시예에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 위한 그래프이다.
- 도 4는 기존의 PFC 정지 프레임과 본 발명의 일 실시예에 따른 PFC 정지프레임을 보여주는 도면이다.
- 도 5는 오픈소스기반 네트워크 기능 시뮬레이터를 설명하기 위한 개략적인 도면이다.
- 도 6은 실험예에서 사용한 우선순위 별 플로우 전송 속도 및 크기를 나타낸다.
- 도 7은 기존의 PFC와 본 발명의 일 실시예에 따른 PFC의 실험 결과값을 보여준다.

발명을 실시하기 위한 구체적인 내용

- [0012] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시 예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시 예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시 예들은 다양한 형태들로 실시될 수 있으며 본 명세서에 설명된 실시 예들에 한정되지 않는다.
- [0013] 본 발명의 개념에 따른 실시 예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시 예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시 예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0014] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.
- [0015] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0016] 본 명세서에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야

한다.

- [0017] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0018] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시 예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시 예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0020] 우선, 본 발명에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 전에 용어를 간략하게 정의하기로 한다.
- [0021] 순위 기반 흐름 제어(priority-based flow control, PFC)는 무손실 이더넷(lossless Ethernet)을 지원하기 위한 대표적인 흐름 제어 기법이다. 구체적으로, PFC는 스위치의 복수의 서비스 클래스(Class of Service, CoS) 타입에 대한 입력/출력 큐의 사용량 모니터링을 통하여 큐의 혼잡 여부를 확인하고, 큐의 사용량이 관리자가 정한 지정량보다 클 경우, 이전 노드 스위치에 정지 프레임(Pause Frame)을 전달하여 혼잡이 생긴 큐에 대한 패킷을 잠시 정지시키는 방식으로 큐 오버플로(Overflow)를 방지하는 기법이다.
- [0022] 기존 PFC기법에서는 지원 가능한 CoS 타입의 종류 제한적이다. PFC는 IEEE 802.1p에서 정의한 8개의 CoS 타입에 대한 트래픽에 대하여 흐름 제어를 진행하는데 이는 데이터 센터의 실시간으로 발생하는 대규모의 플로우를 매핑(Mapping)하고 처리하기에는 제한적이다. 더불어, 데이터 센터의 각 트래픽 플로우를 분석 결과, 저지연을 요구하는 쿼리 트래픽이 파일 전송 트래픽으로 인하여 정체가 발생한다.
- [0024] 이하, 첨부된 도면을 참조하여, 본 발명의 실시예들에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기로 한다. 상술하는 트래픽 제어 방법은 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치, 예컨대 상기 프로세서에 의해 수행될 수 있고, 상기 컴퓨팅 장치는 스위칭 장치를 의미할 수 있다.
- [0025] 도 1은 본 발명의 일 실시예에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 위한 순서도이고, 도 2는 본 발명의 일 실시예에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 블록도이다.
- [0026] 도 1 및 도 2를 참조하면, 같은 CoS 출력 큐에 저장된 패킷을 제어하기 위하여 각각의 패킷에 우선순위 값을 표기하며 이를 랭크(Rank)라고 칭한다(S100). 우선순위가 높을수록 패킷은 낮은 랭크 값을 가진다. 이때, 패킷의 우선순위를 스위치의 테이블 룩업(Table Lookup) 기능을 통하여 확인할 수 있다. 예를 들어, IP 패킷일 경우, 서비스 타입을 나타내는 DSCP(Differentiated Service Code Point) 필드를 조회하여 해당 패킷의 우선순위를 확인할 수 있다.
- [0027] 다음으로, 스위치의 각 CoS 큐의 사용량을 모니터링하여 혼잡 여부를 확인한다(S200).
- [0028] 큐의 사용량이 한계값을 넘을 경우, 우선 상위 스위치 노드에 정지 프레임을 전송하여 일정 기간 동안 낮은 우선순위(높은 랭크) 트래픽에 대한 정지를 요청한다. 해당 시점, 즉 정지 프레임이 해당 스위치 노드에 전송된 시점부터는 상위 스위치로부터 높은 우선순위(낮은 랭크)의 트래픽만을 수신할 수 있다(S300).
- [0029] 한편, 높은 우선순위 트래픽으로 인하여 큐의 정체가 발생할 경우, 상위 스위치에 정지 프레임을 전송하여 모든 트래픽을 정지할 수 있다(S400).
- [0030] 큐의 사용량이 특정 한계값 이하로 내려갈 경우 상위 스위치의 트래픽에 대한 정지를 해제한다. 이러한 방법으로 큐 오버플로우 문제를 해결하여 최종적으로 무손실 이더넷을 보장한다.
- [0031] 일 실시예에 따르면, 각 CoS 출력 큐는 PIFO(Push-in First-out) 큐를 사용한다. PIFO 큐는 높은 우선순위를 가지는 패킷에 대하여 낮은 큐 지연을 보장할 수 있다. 구체적으로, PIFO 큐는 패킷이 큐에 저장되는 시점에 패킷의 랭크 값과 큐에 저장된 패킷들의 랭크값 비교를 통하여 저장 위치를 선정한다. 따라서, 높은 우선순위를 가지는 패킷은 상대적으로 낮은 우선순위를 가지는 패킷보다 큐의 앞에 저장되며 먼저 출력된다.

- [0033] 도 3은 본 발명의 실시예들에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 설명하기 위한 그래프이다.
- [0034] 도 3을 참조하면, 각 큐의 사용량과 한계값과의 비교를 통하여 현재 큐의 정체 여부를 판단할 수 있다. 복수의 한계값들을 가질 수 있으며 해당 한계값에 도달하면 하기와 같은 단계들이 수행될 수 있다. 또한, 본 명세서에 기재된 복수의 한계값들은 사용자나 관리자 등에 의하여 미리 결정된 값을 의미할 수 있으며, 상황에 따라 그 값은 변화할 수 있다.
- [0035] 1) XON: 현재 큐가 정체되지 않는 상태를 표현하는 한계값이며, 큐 사용량이 XON 이하로 내려갈 경우, 상위 스위치가 정지 상태이면 해당 CoS 큐에 대하여 T 값을 0을 가지는 정지 프레임을 상위 스위치에 전달한다. 실시예에 따라, 정지 프레임은 모든 상위 스위치에 전달될 수 있다.
- [0036] 2) XOFF_Low: 현재 큐가 모든 우선순위의 트래픽에 의하여 정체가 발생됨을 나타내는 한계값이다. 큐 사용량의 XOFF_Low에 도착할 경우, 상위 스위치에 정지 프레임을 보내어 XOFF_LOW에 설정한 정지 랭크값을 전달한다.
- [0037] 3) XOFF_ALL: 현재 큐가 높은 우선순위의 패킷에 의하여 정체가 발생함을 나타내는 한계값이다. 큐 사용량의 XOFF_ALL에 도착하면, 모든 트래픽에 대한 정지를 요청하는 프레임을 상위 스위치에 전달한다.
- [0038] 본 실시예에서는 설명의 편의를 위하여 도 3에서 1개의 XOFF_Low 한계값만 도식화하였으나, 실제 사용에서는 다중 XOFF_Low 한계값을 설정하여 다중 레벨(Multi-Level)에서의 트래픽 관리가 가능하다.
- [0039] 본 발명의 일 실시예에 따르면, 정지 프레임은 {C, T, R}의 집합으로 표기할 수 있다. C는 CoS 값(즉, 해당 큐에 대응하는 CoS 타입을 식별할 수 있는 정보)을 나타내고, T는 정지 기간을 말하며 단위는 512bit 전송에 필요한 시간으로 “quanta”로 표현될 수 있다. R은 정지할 트래픽의 우선순위 최고 랭크 값을 말한다. 예를 들어 {C=3, T=10, R=3}은 “CoS 3인 큐에 대하여, 랭크 값이 3보다 높은 모든 트래픽을 10 quanta 시간 동안 정지”를 요청하는 메시지를 나타낸다. 추가로 T 값이 0일 경우는 “정지 해제”를 의미하고 R 값이 0이면 “전체 트래픽 정지”를 의미한다.
- [0041] 도 4는 기존의 PFC 정지 프레임과 본 발명의 일 실시예에 따른 PFC 정지 프레임을 보여주는 도면이다.
- [0042] 도 4를 참조하면, 본 발명의 일 실시예에 따른 정지 프레임은 64 바이트이며 앞 34 바이트는 기존 PFC와 같은 구조를 가진다. 자세한 설명은 다음과 같다. 정지 프레임 앞 14 바이트는 이더넷 헤더이고 DMAC 필드와 EtherType 필드는 각각 멀티캐스트 주소 “0x01:80:C2:00:00:01”와 “0x8808”을 가진다. 2 바이트의 OpCode는 값 “0x0101”을 가지며 이는 PFC 계열의 정지 프레임을 의미한다. 2 바이트 길이의 Priority Enabled Vector은 내부적으로 1 바이트를 사용하여 8-bit 벡터로 8 개의 CoS 값을 표시한다. 16 바이트의 Time 필드는 각 CoS에 해당되는 정지 기간을 나타낸다. 본 발명의 일 실시예에 따른 PFC에서는 추가로 16바이트의 Pause Rank 필드를 사용하여 각 CoS 트래픽의 정지할 랭크 최대값을 표기한다. 이와 달리 기존의 PFC에서는 해당 Pause Rank 필드 자리에는 0 값으로 패딩(Padding)하여 사용한다. 따라서 이론상 PFC 기능을 지원하는 스위치와 본 발명의 일 실시예에 따른 PFC를 지원하는 스위치로 구성된 토폴로지에서도 무손실 이더넷 보장이 가능하다.
- [0043] 이로써, 큐 혼잡을 탐지 시, 낮은 우선순위 패킷을 우선적으로 정지하여, 높은 우선순위 패킷은 낮은 우선순위 패킷에 비해 상대적으로 낮은 큐 지연 및 높은 대역폭을 가진다.
- [0044] 또한, 낮은 우선순위 트래픽에 대한 정지가 발생 시, 새로 생성된 높은 우선순위 트래픽에 대해서는 정상적으로 출력한다. 그리고 높은 우선순위 트래픽에 인하여 큐 혼잡이 발생 시, 전체 트래픽을 정지하는데, 높은 우선순위 트래픽의 대역폭이 해당 큐의 출력 속도보다 낮을 경우, 높은 우선순위 트래픽의 입력 대역폭과 같은 출력을 보장한다.
- [0045] 같은 CoS 값을 가지는 높은 우선순위 트래픽과 낮은 우선순위 트래픽에 대하여, 높은 우선순위 트래픽이 상대적으로 낮은 큐 지연 및 높은 대역폭을 가질 수 있다. 최종적으로 높은 우선순위 트래픽이 더 낮은 FCT(Flow Completion Time)를 가질 수 있다.
- [0046] 한편, 다중 지정량에 대한 설정을 통하여, 능동적으로 큐 혼잡제어가 가능하다.

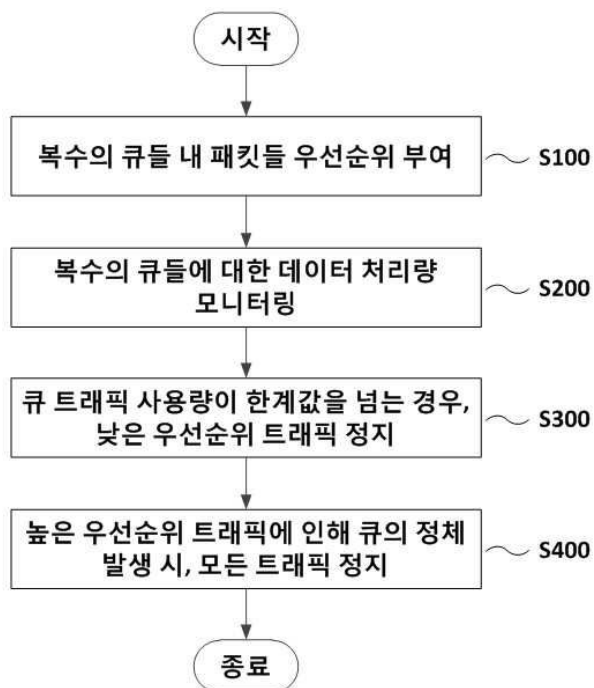
- [0048] 이하, 본 발명의 일 실시예에 따른 이더넷 기반 네트워크 내 트래픽 제어 방법을 이용한 네트워크에 대하여 실험예를 통해 설명하기로 한다. 본 실험예는 오픈소스기반 네트워크 기능 시뮬레이터인 NS-3에서 실험을 진행하였다.
- [0049] 도 5는 오픈소스 기반 네트워크 기능 시뮬레이터를 설명하기 위한 개략적인 도면이다.
- [0050] 도 5를 참조하면, 4 대의 송신 노드(Sender)와 1 대의 수신 노드(Receiver) 각각 1Gbps 링크로 중앙 스위치에 연결되었다. 각각의 송신 노드는 같은 16 개의 플로우(또는 애플리케이션)가 있으며 동시다발적으로 수신 노드에 패킷을 전달하며 모든 플로우의 CoS 값은 같다. 각 플로우의 송신 속도 및 크기는 각각 평균 값 62.5Mbps와 10MB를 가지는 정규분포를 따르도록 설정하였으며, 실험에서는 플로우의 우선순위 범위는 1부터 16까지이고 균일 분포를 따르도록 설정하였다.
- [0051] 도 6은 실험예에서 사용한 우선순위 별 플로우 전송 속도 및 크기를 나타낸다. 도 6을 참조하면, 우선순위를 기준으로 집결한 플로우의 송신속도 및 크기를 보여준다. 추가로 큐의 크기는 1000 패킷이고 XON, XOFF_Low, XOFF_ALL의 한계값은 각각 128, 400, 700으로 설정하였다.
- [0052] 도 7은 기존의 PFC와 본 발명의 일 실시예 따른 PFC의 실험 결과값을 보여준다.
- [0053] 실험예(4)는 XOFF_Low 한계점에 해당하는 정지 랭크 값이 4임을 나타낸다. 실험예(4) 시나리오에서는 높은 우선순위 패킷들의 송신 속도 합은 큐 출력 속도보다 낮아서 XOFF_ALL 한계값을 넘지 못했다.
- [0054] 실험예(7)은 XOFF_Low 한계점에 해당하는 정지 랭크 값이 7임을 나타낸다. 실험예(7) 시나리오에서는 높은 우선순위 패킷들의 송신 속도 합은 큐 출력 속도보다 높으므로 XOFF_ALL 한계점으로 인한 흐름 제어가 발생하였다.
- [0055] 실험예(4) 및 실험예(7)은 기존 PFC에 비해 높은 우선순위 플로우에 대한 FCT(Flow Completion Time)을 각각 70%, 42.8% 단축시켰다. 비록 낮은 우선순위 플로우에 대한 FCT는 5.6%, 11.3%의 증가가 발생하였지만 전체 플로우에 대한 평균 FCT값을 8.26%, 9.23% 단축 시켰다. 또한, 각각의 시나리오에서 발생한 정지 프레임 카운터(Counter)값은 도 7에서 확인 가능하다.
- [0056] 추가로 실험예(4)의 실험데이터를 분석 결과, 높은 우선순위 플로우들에 대해서는 0에 가까운 큐 지연을 확인할 수 있었다. 이는 데이터 센터의 실시간으로 발생하는 대규모의 플로우들로 정체가 발생시, 본 발명에 따른 PFC를 사용하면 높은 우선순위에 대한 출력 속도를 보장할 수 있다.
- [0058] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시 예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPA(Field Programmable array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(Operation System, OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.
- [0059] 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(Code), 명령(Instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(Collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0060] 실시 예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판

독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시 예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(Magnetic Media), CD-ROM, DVD와 같은 광기록 매체(Optical Media), 플롭티컬 디스크(Floptical Disk)와 같은 자기-광 매체(Magneto-optical Media), 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시 예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

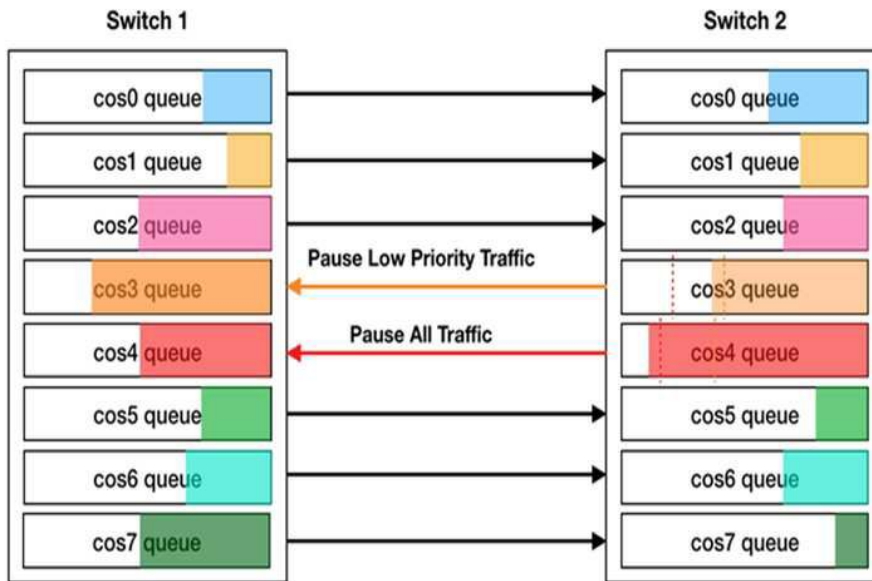
[0061] 본 발명은 도면에 도시된 실시 예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

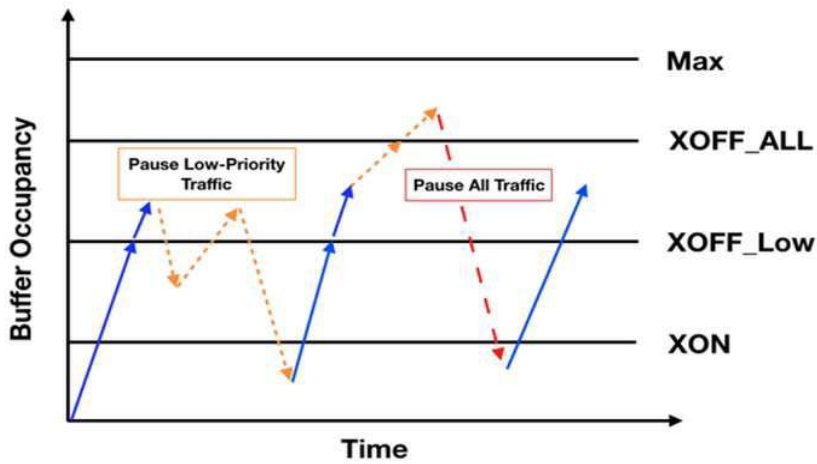
도면1



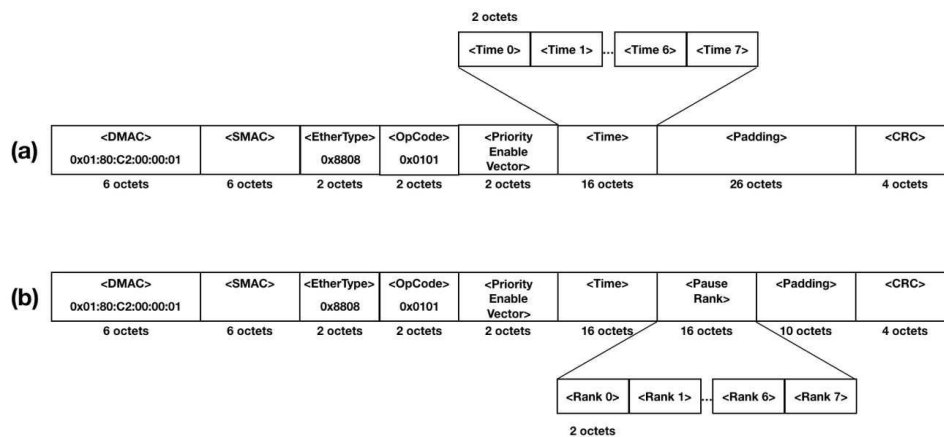
도면2



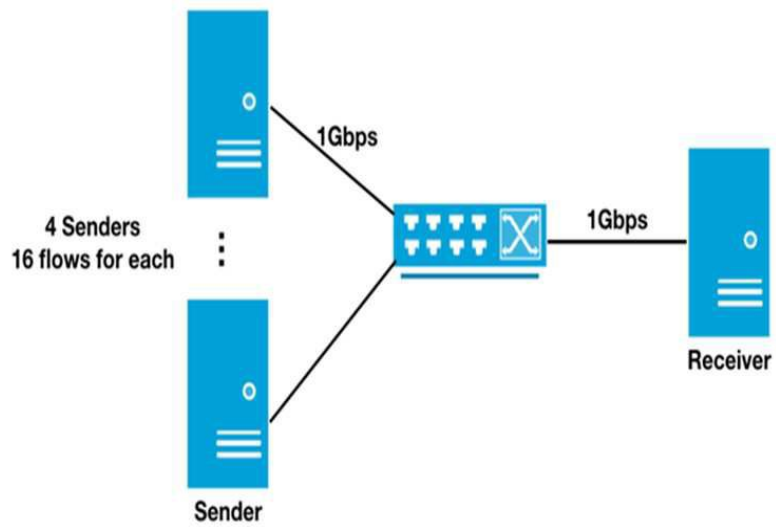
도면3



도면4



도면5



도면6

Priority	Flow Size(MB)	Flow Rate(Mbps)	Agg Flow Rate(Mbps)
P1	41.53344	283	283
P2	85.325824	549	832
P3	18.021888	144	976
P4	60.902912	400	1376
P5	53.576704	314	1690
P6	22.045696	152	1842
P7	80.323584	547	2389
P8	12.364288	65	2454
P9	36.937216	219	2673
P10	6.303232	64	2737
P11	44.990464	326	3063
P12	22.188544	118	3181
P13	42.432512	216	3397
P14	36.586496	277	3674
P15	19.324416	147	3821
P16	59.237376	446	4267
Total	642.094592	4267	4267

도면7

	Simulation 1: Pause Rank 4		Simulation 2: Pause Rank 7	
	High Priority	Low Priority	High Priority	Low Priority
Flow Size(MB)	144.881152	497.21344	281.406464	360.688128
Flow Rate(Mbps)	976	3291	1842	2425
	PFC	G-PFC	PFC	G-PFC
Tail FCT for High Priority	4505.72391	1340.032683	4950.967452	2773.986618
Tail FCT for Low Priority	5177.697648	5263.411752	5177.697648	5428.934302
Avg FCT	4484.115861	4113.52619	4484.115861	4069.867363
XOFF_ALL Count	1320	0	1320	549
XOFF_Low Count	0	2090		1548