

고속 라우터의 성능 향상을 위한 TCP MAC+ 블록 기능부
구현 연구

A Study of TCP MAC+ Block Functional Part upgrading
High Speed Router Performance

우 상 철(정보통신과)

Sang-Choel Woo (Dept. of Information & Communication)

Key Words : MAC(Media Access Cntrol),TAC(TCP/IP Acceleration Card)

ABSTRACT: In this research, we consider one topic for improving performance of high-speed routers. It is to implement the receiving part of MAC+ which can do the TCP checksum operation on the silicon. For the MAC+, it is important to acquire high-speed processing technology for protocols TCP/IP and acquire its ASIC technology. For the MAC+, we implement MAC controller logic extract TCP/IP Acceleration module from TCP/IP TAC which has been developed last year.

1. 서론

Unix 및 Window NT 등은 커널 모드와 사용자 모드로 분리된 이중구조를 가진다. 이러한 구조 속에서 응용 프로세스와 커널은 각각 독자적인 메모리를 사용하기 때문에 응용 프로세스가 소켓함수 호출을 이용하여 커널내에 위치한 TCP/IP로 접근할 때는 데이터 블록의 복사가 필수적이다[1][2]. 데이터를 송신할 때, 응용 프로세스가 생성한 데이터 블록은 소켓 시스템 콜에 의해 커널의 mbuf로 복사된다. 이후, 복사된 데이터 블록의 TCP/IP MAC 헤더 부분을 위한 mbuf가 새로이 할당된다. TCP/IP MAC 헤더 부분에 대한 구성이 완료되면 데이터 부분의 mbuf와 헤더 부분의 mbuf가 연결되어 LAN과 공유하는 송신 메모리에 복사 전달된 다음 송신된다. 수신되는 데이터 프레임의 바이트열은 우선 수신 메모리에 저장되고 수신이 완료되면 LAN카드가 인터럽트를 발행한다. 이 인터럽트에 의한 처리로 커널이 준비한 수신용 mbuf에 수신 데이터를 복사, 이동한다. 수신 프레임에 대한 헤더 검사가 TCP/IP에 의해 완료되면 운영체제의 스케줄링에 의해 mbuf에 있는 데이터가 사용자모드의 메모리에 복사 전달된다[3].

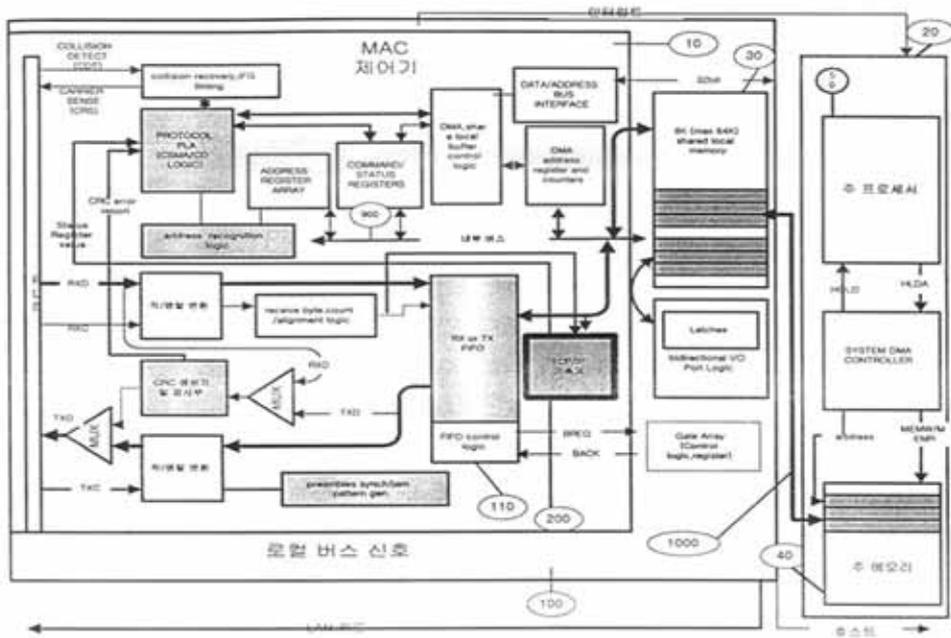
이러한 커널모드와 사용자 모드로 분리된 이중 구조 방식을 사용하면, 응용 프로세스는 소켓함수 호출을 사용하여 프로토콜 동작과 무관하게 송수신을 수행할 수

있다는 장점이 있다. 그러나 TCP/IP가 커널속에 있기 때문에 두 번의 복사동작이 수행되어 비 효율적이다. 이외에도 TCP나 IP의 체크섬 계산시 소프트웨어 방식으로 수행됨으로, 데이터의 프레임 길이가 길수록 CPU 사이클을 더 많이 소비하는 문제점이 발생한다. 하드웨어로 처리되는 프로토콜은 CPU 사이클과 무관할 뿐만 아니라 기계적인 성질로 인하여 고속 처리가 가능하다. 따라서 TCP/IP가 하드웨어로 구현된다면 그 성능향상은 소프트웨어 방식에서 실현되는 것 보다 훨씬 더 클 것으로 기대된다[4].

TCP/IP를 고속화하는 가장 근본적인 방안은 결국 다른 응용 프로그램을 처리하기 위해 바쁜 Host CPU로부터 TCP/IP처리를 독립시킴으로서 Host CPU가 다른 일로 바쁠때에도 TCP/IP 처리 효율은 일정치 이상 유지 하도록 하는 것이다. 따라서 프로토콜 대부분을 소프트웨어 방식으로 처리하되 체크섬이나 헤더 검사 같은 기계적 특성이 있는 영역을 하드웨어로 구현하는 하이브리드 방식이 실현될 가능성이 높다. 본 연구에서는 하이브리드 방식으로 TCP/IP 모듈을 장착하는 MAC+를 구현하고 그 성능을 고찰한다.

2. MAC+의 구성

MAC+는 고속 라우터를 구현하기 위하여 MAC Controller내부에 TCP 체크섬 과정을 하드웨어로 처리하는 TCP/IP 가속기를 부착함으로써 TCP 체크섬 과정을 소프트웨어적으로 처리하여 호스트 CPU에 걸리는 부하를 제거하여 고속의 처리를 이루는 off-loading 방식이다. MAC+에서는 송수신 FIFO에서 로컬 메모리로 수신 프레임을 저장하면서 동시에 헤더 부분을 체크하여 TCP/IP의 체크섬 계산 과정을 수행한다. 그리고 체크섬 에러나 FCS에러가 발견될 경우 로컬 메모리에서 해당 패킷을 지워 버려서 상위 프로토콜로 전달하지 않도록 한다. 상위 계층으로 전달된 패킷은 체크섬 계산이 끝난 패킷임으로 소프트웨어 TCP/IP에서는 checksum()함수를 명시함으로 체크섬 계산에 드는 시간을 단축시킬 수 있다. MAC+ 보드의 기능적인 블록 구조는 다음 <그림 1>과 같다.



<그림 1> A Block Diagram of MAC+

MAC+의 기능적인 블록 내부를 보면 크게 다음과 같이 분류할 수 있다.

(가) MAC Controller 수신부

- SFD(Starting Frame Delimiter) 검사부
- CRC(Cyclic Redundancy Check) 검사부
- Serial to Parallel 32 비트 변환부
- 주소 검사부
- 수신 버퍼

(나) TCP/IP 가속기부

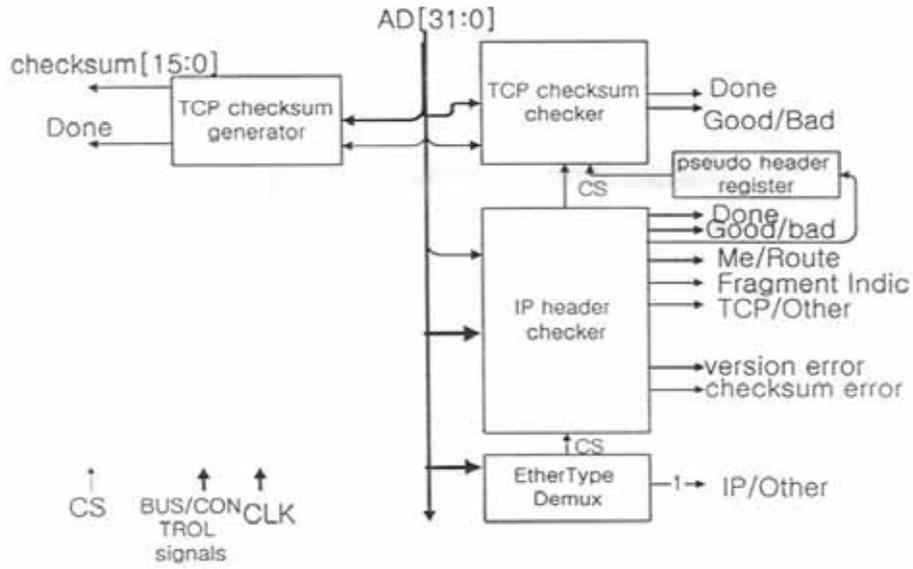
- 전방 처리기
- IP 기능부
- TCP 기능부

MAC+의 동작을 살펴보면, 먼저 PLS(Physical Layer Signaling)를 거친 패킷이 쉬프트 레지스터를 통해 직/병렬 변환되는 동안 CRC가 Preamble 비트열로 시작하는 프레임에서 SFD를 감지한다. 수신 FIFO로 들어가는 동안 동시에 주소 감지로

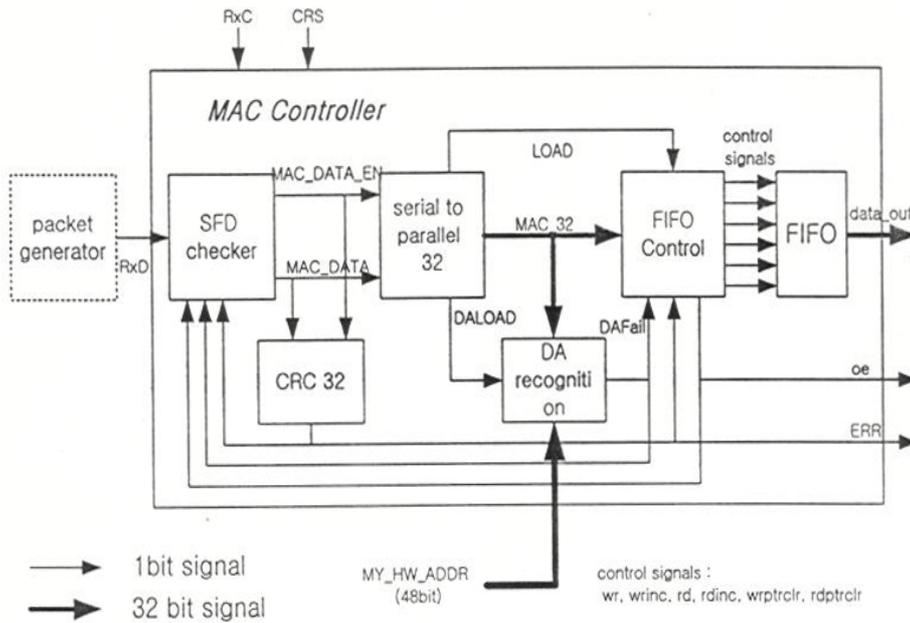
직에서는 6바이트의 목적지 주소가 유효한지 검사한다. 만약 주소가 유효하지 않을 경우 PLA에 신호를 보내서 FIFO에서 해당 데이터를 삭제토록 한다. 주소 유효성 검사에서 이상이 없을 경우 FIFO에 저장되어 있던 데이터는 로컬 메모리에 저장되면서 동시에 TCP/IP 가속기에 의해 체크섬 검사를 받는다.

FCS는 MAC프레임의 가장 끝에 32비트 크기를 가지고 있으며 FCS필드를 제외한 MAC 프레임의 CRC 체크값을 저장하고 있다. 그리고 CRS(Carrier Sense)는 데이터 필드의 끝이 아닌 FCS필드의 끝에서 “0”으로 떨어짐으로 그때까지의 CRC 체크값은 FCS필드를 포함한 계산값이 된다. 따라서 CRC 검사부에서는 들어오는 데이터를 32 비트 크기를 가지는 shift register에 넣어서 32 클럭 동안 지연시킨 후 CRC검사를 하면 CRS가 “0”이 되었을 때의 CRC 체크 계산값은 FCS 필드를 제외한 영역의 계산값이 되며, 그때 32클럭 지연용으로 만든 shift register에 있는 값이 FCS 영역의 값이 된다. 다음 <그림 2>는 MAC Controller 수신부의 구성이다. 그리고 다음 <그림 3>은 TCP/IP 블록의 구성도이다. MAC Controller 수신부는 PLS(Physical Layer Signaling)에서 preamble를 감지하여 동기를 맞춘 RxC(Receive Clock)와 CRS(Carrier Sense)신호에 의해서 동작한다. PHY를 거쳐서 MAC Controller로 전달되는 데이터는 preamble와 SFD가 제거되지 않은 데이터이다. 따라서 SFD Checker에서 이를 감지하고 제거해 준다. Preamble에서는 3-7바이트의 0과 1이 반복되는 형태이고 SFD는 preamble 뒤에 1이 한 바이트 반복된다. PHY에 전달된 데이터는 preamble이 감지된 상태임으로 SFD Checker에서는 SFD만을 검사하여 그 뒤의 데이터만을 다음 로직으로 전달해 준다.

CRC32 로직은 수신된 MAC 프레임의 FCS(Frame Check Sequence)를 검사한다. MAC 프레임은 송신될 때 CRC를 행하여 프레임의 마지막 부분에 4바이트의 CRC값을 추가한다. 수신측은 MAC 프레임이 수신되기 시작할 때부터 CRC 검사를 수행하여 마지막으로 FCS값과 CRC값을 비교한다. 값의 차이가 틀릴 경우 결과 레지스터에 그 값을 써서 상위 모듈이 해당되는 패킷을 폐기하도록 한다. MAC 프레임은 Serial/Parallel 32bit 로직을 거쳐서 32비트 병렬구조로 변환시킨 뒤 32x16 비트의 크기를 가지는 임시버퍼와 address recognition로직으로 동시에 전달된다.

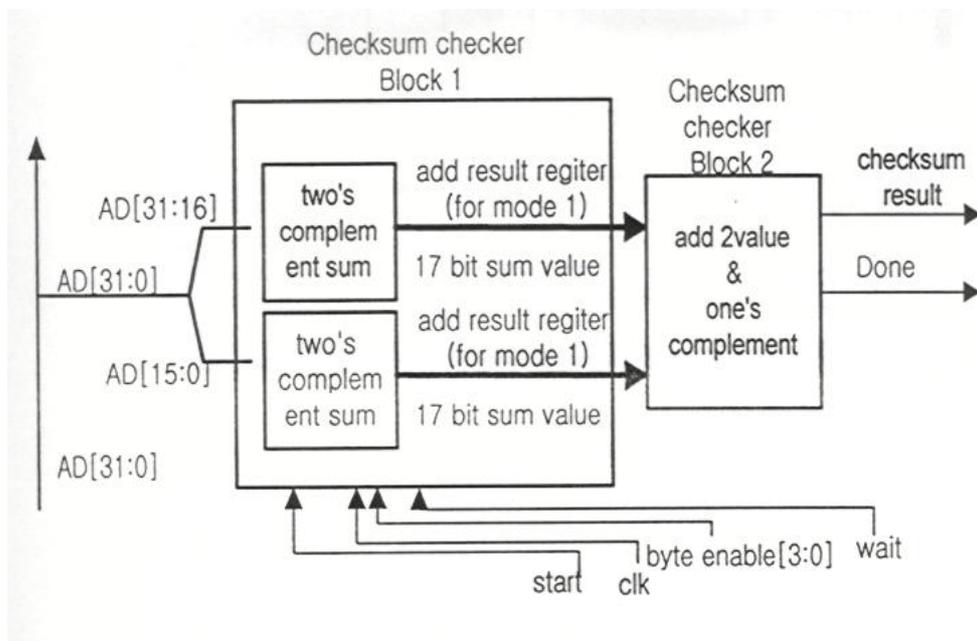


<그림 2> A Diagram of MAC Controller Receiver



<그림 3> A Block Diagram of TCP/IP

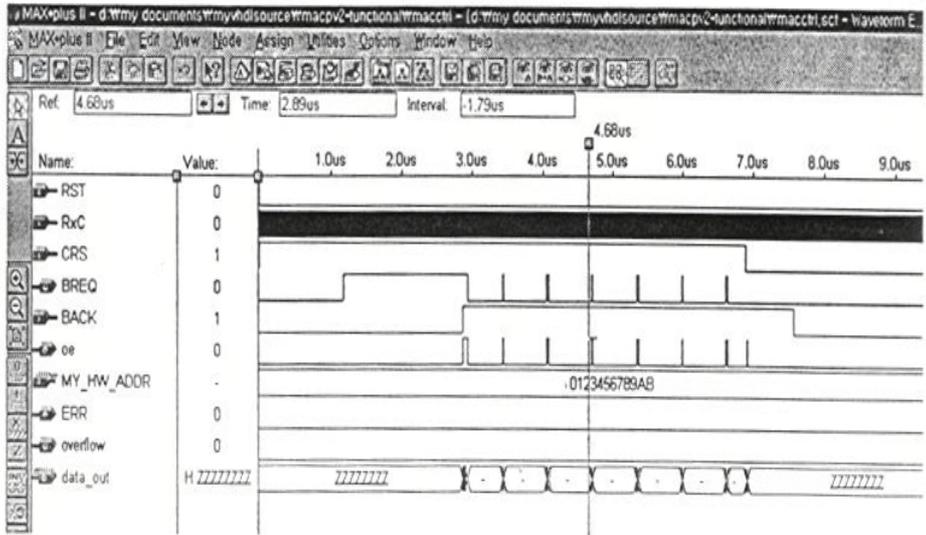
TCP/IP용 체크섬 검사기의 구조는 다음 <그림 4>와 같다. IP가 검출되면 IP헤더 부분의 체크섬이 시작된다. 체크섬은 32 비트의 데이터 버스를 16비트씩 분할하여 데이터 버스의 하위 16 비트 덧셈 블록과 상위 16 비트 덧셈 블록에서 동시에 수행하도록 하였다. 이렇게 하는 가장 큰 이유는 33MHz 클럭으로 이동되는 32 비트 데이터를 16 비트로 분할하여 더하는 과정에서 1 클럭에 32 비트 데이터 버스의 값에 대한 연산을 수행토록 하여 체크섬 과정의 지연을 줄이기 위함이다. 각각 독립적으로 수행된 상위, 하위 16 비트 데이터에 체크섬 결과는 이후 헤더에 대한 연산이 종료된 후 서로 더해지게 된다. 상위, 하위 데이터들의 덧셈 과정에서 발생한 캐리(carry)는 최종 블록으로 넘겨지도록 하였다. 또한, 상위 덧셈부는 동작 신호를 알리는 신호를 받으면 클럭에 따라 1씩 증가하는 카운터 동작을 하면서 카운터 값이 IP의 헤더 값과 같아지면 그 동작을 정지시키고 최종 처리블럭을 동작시킨다.



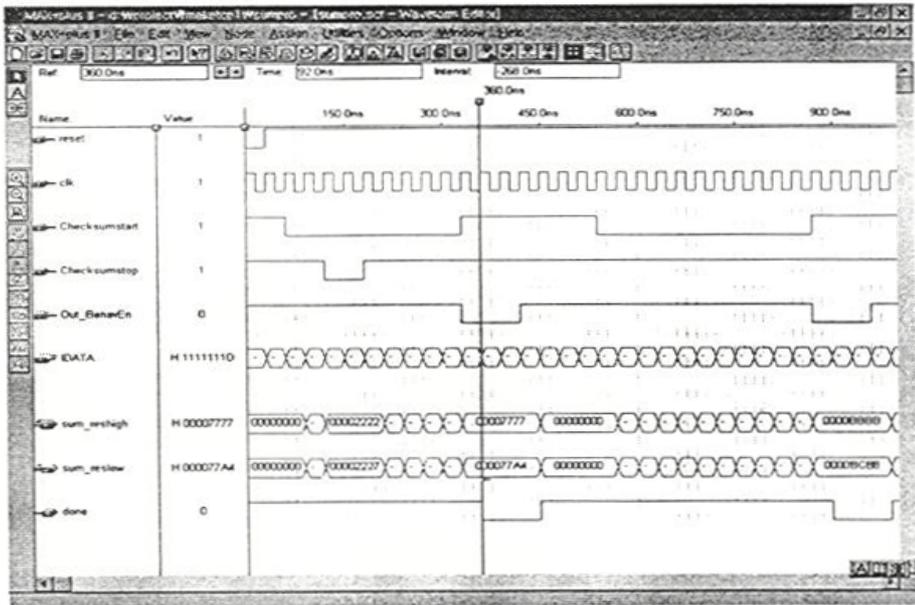
<그림 4> The Format of TCP/IP Checksum

3. TCP/IP MAC+ 기능부 구현 및 시뮬레이션

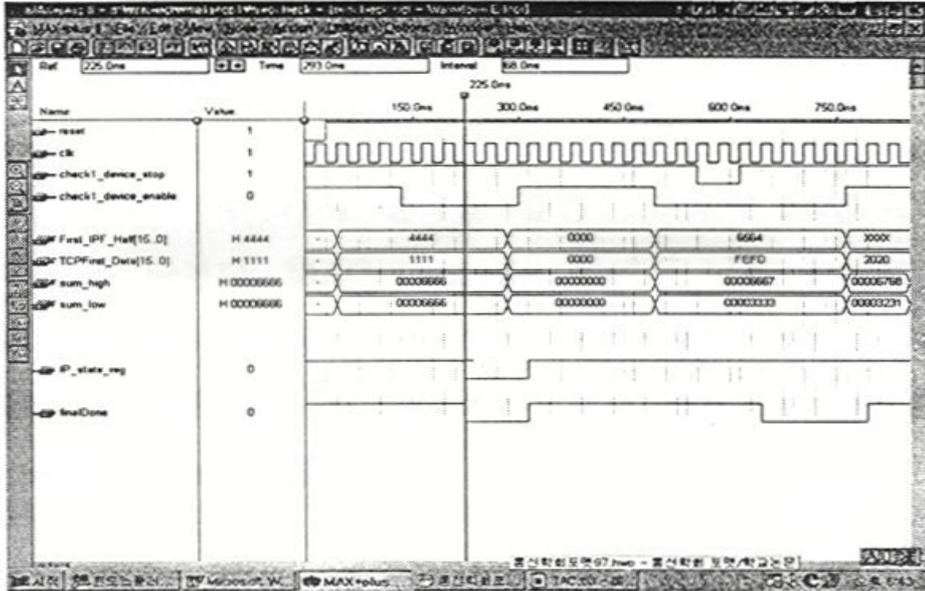
TCP/IP MAC+ 기능부의 설계를 통한 MAC Controller 수신부의 VHDL 코딩의 시뮬레이션 결과는 다음 <그림 5>와 같다. 그리고 이더넷 유형에 따른 IP 데이터



<그림 6> IP Datagram Part



<그림 7> 16 bit Functional Part



<그림 8> Final Checksum Part

4. 결론

본 연구에서 제안된 MAC+는 인터넷 망에서 고속 라우터를 구현하기 위하여 MAC Controller내부에 TCP 체크섬 과정을 하드웨어로 처리하는 TCP/IP 가속기를 부착함으로써 TCP 체크섬 과정을 소프트웨어적으로 처리하여 호스트 CPU에 걸리는 부하를 제거하여 고속의 처리를 이루는 off-loading 방식이다. 그리고 MAC+ 방식 연구의 경우 고속 통신망 발전에 의한 단말에서의 고속 TCP/IP 처리의 필요성 및 TCP/IP ASIC 개발의 기반 기술 확보에 그 중요성이 있다. MAC+의 경우 이미 확보한 기술인 TAC에서 패킷 수신 모듈을 분리하였고 MAC Controller 수신부를 설계하였다. 또한, FPGA를 이용하여 MAC의 처리속도를 향상시키기 위하여 체크섬 부분을 모든 VHDL 코드를 합성 가능한 코드로 수정하였다.

참고문헌

- [1] A Gallatin, "TCP/IP : at Near-Gigabit Speeds", JCCI'98, pp127-132
- [2] Jau-Hsiung and Chi-Wen Chen, "On Performance Measurements of TCP/IP and its

Device Driver," 17th LCN, 1992.

[3] Clark, Jacobson, Romkey, and Salwen, "An analysis of TCP processing overhead" IEEE Comm. Mag., June, 1989, pp. 23-29

[4] H. Kim et al., "Performance analysis and feasibility study of a parallelized TCP/IP",
JCCI' 97, pp. 443-446