

## QorIQ매물형처리기의 자료경로가속구성방식과 LPM 경로표를 리용한 고속망경로기의 구성방법

김수일, 박진혁

경애하는 김정은동지께서는 다음과 같이 말씀하시였다.

《정보통신의 현대화에서 나서는 중심과업은 전신, 전화의 자동화, 수자화, 빛섬유케블화, 컴퓨터화의 성과를 공고발전시키면서 완전IP화, 고정통신과 이동통신의 통합화, 통신망의 광대역화를 전면적으로 실현하는것입니다.》

IP부분망사이에 자료를 교환할 때에는 송신측에서 IP패킷을 목적부분망의 경로기로 보내고 그 경로기가 목적지로 패킷을 전송한다. 이때 경로기의 기본동작은 포구검출과 전송으로서 포구검출의 목적은 목적지의 최단경로를 판정하여 경로를 선택하는것이다.

한편 서로 다른 경로선택규약과 경로선택산법으로 최단경로를 판정하기 위하여 여러 가지 경로표가 리용되는데 합리적인 경로표의 구성은 자료교환의 속도와 경로화과정의 효율을 높일수 있게 한다.

현재 경로기들에서는 2진나무나 Trie나무를 리용한 경로표를 많이 리용하고있다. 그러나 이러한 구성방식은 경로표의 수가 늘어나면 목적지포구의 탐색시간이 증가하며 기억기소모도 커진다.[1]

ASIC형식의 대규모집적소자로 구성된 소규모의 보통경로기에서는 기억기소모를 줄이기 위하여 LPM(Longest Prefix Match)표를 리용한 경로화를 리용한다.[2]

또한 고속망경로기에서는 IP패킷의 처리를 하드웨어적으로 진행하거나 응용층에서 직접 처리하여 자료처리속도를 높일수 있게 구성되여있다.

QorIQ계열의 매물형처리기에서는 자료통로가속구성방식(DPAA)기능을 리용하여 IP패킷의 분류와 려파, 분배를 하드웨어적으로 진행함으로써 망응용프로그램의 성능을 높일수 있으며 처리기의 부하를 줄일수 있다.

본문에서는 QorIQ계열 T2080매물형처리기의 DPAA기능을 리용하는 경로기에서 LPM 경로화표를 적용하여 경로기의 성능을 높이기 위한 방법을 제안하고 2진나무나 Trie나무 경로표를 리용한 방법과 비교하여 그 효과성을 검증하였다.

### 1. QorIQ계열매물형처리기에서 DPAA체계를 리용한 경로기의 구성

QorIQ계열처리기에서는 자료통로가속화구성방식(DPAA)이라고 하는 체계를 리용하여 다중처리기핵심부들이 망대면부들과 자료처리가속기들에 대한 공유를 간단하게 진행하는 하부구조를 제공한다.

DPAA는 동일한 패킷처리와 순서화의 요구를 가진 연속적인 패킷들로 구성된 망패킷흐름을 단위로 처리를 진행한다.

일반적으로 어떤 통신대면부에 도착한 패킷들은 패킷분석과 분류방법 실행으로 조종평면자료흐름에 따라 해당한 처리기핵심부들과 하드웨어가속기들로 분배된다. 조종프로그램은 패킷흐름안의 패킷들을 어떻게 처리하는가를 표현하는 자료구조체들을 갱

신 또는 창조하여 하드웨어적으로 이러한 흐름들의 처리를 진행한다. 이 자원들은 DPAA의 대기렬관리자를 리용한 자료들의 대기렬조작에 의하여 진행된다.

### 1) DPAA의 기능

#### ① 파के트분산처리와 대기렬자료관리 및 파케트혼잡관리기능

여기에는 자료완충기관리, 대기렬관리, 파케트의 대기렬분배, 파케트흐름조종방책관리기능들이 있다.

#### ② 자료내용처리의 하드웨어적인 가속화처리기능

이 기능에서는 자료의 압축 및 해제, 암호화 및 복호화, 기호열검색, RapidIO대면부의 통보관리와 같은 기능들을 하드웨어적으로 실현한다.

### 2) 자료통로에서 망파케트처리단계

자료통로로 들어오는 망파케트는 다음의 단계를 거쳐 처리된다.

① 자료경로처리는 망파케트가 망대면부에 도착했을 때 시작된다. 파케트가 외부망대면부가 아닌 다른 주변장치모선을 통하여 도착하는 경우에도 파케트처리흐름은 같다.

② 프레임관리자(FMan: Frame Manager)는 망파케트를 수신한 후 그것을 저장하기 위해 완충기관리자(BMan: Buffer Manager)하드웨어에 완충기할당을 요청하며 거기에 파케트들을 저장한다. BMan은 완충기묶음을 관리하며 매 완충기묶음은 소프트웨어적으로 할당한다. 이때 FMan은 제일 적당한 완충기묶음으로부터 완충기를 요청할수 있도록 초기화된다. 만일 수신되는 파케트들을 위한 충분히 큰 완충기를 찾을수 없으면 FMan은 파케트를 몇개의 작은 완충기들에 갈라 저장하고 이 완충기들에 대한 분산 및 집합목록을 창조한다.

③ FMan은 파케트를 처리기핵심부에 보내기 위한 초기분류를 진행하거나 특정흐름의 처리를 위해 자료경로처리기들중의 하나로 넘기는 소프트웨어에 의한 추가적인 분류를 진행한다. 망파케트를 특정한 처리기핵심부들로 보내는 방법은 흐름의 QoS속성에 의한 분류방식에 기초한다.

④ 처리기핵심부로 보내기는 FMan이 대기렬관리자(QMan)에서 지정된 프레임대기렬번호(FQID)의 대기렬에로의 넣기명령을 통하여 수행된다. 이때 프레임대기렬(FQ: Frame Queue)은 흐름의 QoS요구에 기초하여 선택될수도 있다.

⑤ QMan은 FQ를 해당한 처리기핵심부에 배당된 통로의 작업대기렬(WQ: Work Queue)에 넣는다. FQ는 이 특정흐름에 속하는 모든 파케트들이 특정의 처리기핵심부에서 처리되도록 구성되거나 이전에 정의된 부하균형방책에 의해 대응되는 핵심부의 대기렬에 있게 된다. 하나의 WQ의 모든 FQ들은 같은 우선권을 가지며 처리기핵심부가 파케트흐름으로부터 받는 자료의량은 사용자에게 의해 구성가능하다.

⑥ 일단 구성된 QMan은 WQ와 그안의 FQ에 대한 일감처리를 진행한다.

⑦ 처리기핵심부는 완충기안의 파케트자료의 변경을 포함하여 파케트자료를 규약에 따라 처리한다.

망파케트가 IPsec처리를 요구하면 처리기핵심부는 IPsec처리와 관련된 FQID를 리용하여 QMan에 되돌려보낸다.

⑧ QMan은 FQ의 다음처리가 암호화 및 복호화처리기(SEC)에서 진행되면 그 FQ를 SEC에 배당된 통로의 WQ에 넣는다. SEC는 그에 배당된 통로에서 WQ로부터 파케트를 꺼낸다.

⑨ SEC는 FQ에서 자료를 꺼내어 처리하며 해당한 마당들을 추가하고 암호화된 자

료를 원래의 완충기 또는 SEC가 BMan에 요청한 새 완충기에 써넣는다. SEC는 패킷자료료를 갱신하고 FQ를 QMan에 되돌려보낸다.

## 2. LPM경로표의 구성

경로기에서 경로표의 합리적인 구성은 목적지경로의 탐색시간을 줄임으로써 경로기의 성능을 높일수 있게 한다.

LPM(Longest Prefix Match)경로표에서는 목적지경로의 검색을 위하여 목적지IP주소의 앞붙이를 리용한다. 경로기에 도착하는 패킷들은 LPM경로화알고리즘에 의하여 목적지 경로로된다.

### 1) LPM경로표의 구성

논문에서는 전통적인 Radix-Trie경로표대신 보다 간단한 LPM경로표를 구성하였다.

LPM경로표에서는 5준위표를 사용한다.

첫째 준위표: IP주소의 첫 16bit에 대응되는 65 536개의 크기를 가진 배열이다.

둘째 준위표: IP주소의 12-15bit에 대응되는 32개의 크기를 가진 배열이다.

셋째 준위표: IP주소의 8-11bit에 대응되는 32개의 크기를 가진 배열이다.

넷째 준위표: IP주소의 4-7bit에 대응되는 32개의 크기를 가진 배열이다.

다섯째 준위표: IP주소의 0-3bit에 대응되는 32개의 크기를 가진 배열이다.

둘째 준위표부터 다섯째 준위표까지는 첫째 준위표의 해당하는 마디가 유효한 값을 가질 때에만 창조된다. 그림 1에 LPM경로표의 구성을 보여주었다.

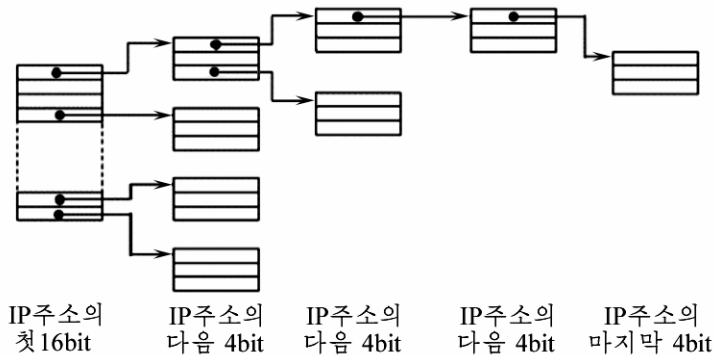


그림 1. LPM경로표의 구성

초기에 IP주소의 첫 16bit에 대응하는 65 536개의 입구를 가지는 첫 준위표만이 창조된다. 다음 경로표에 자료를 추가할 때 IP주소와 망마스크값에 따라 둘째-다섯째 준위 배열들이 창조될수 있으며 목적지IP주소에 해당하는 포구를 검색할 때 최대로 5번 비교를 하게 된다.

### 2) LPM경로표에서 목적지IP에 따르는 포구결정실례

① 판문주소 1.1.1.1, 목적지포구가 1인 A급망의 경로 10.0.0.0/8을 경로표에 추가한다.

10.0.0.0의 첫 16bit는 0x0a00(2560)이다. 망마스크가 8bit로서 첫 준위의 비트길이(16)보다 작으므로 첫 준위표에서 2 560(0x0a00)부터 2 815(0x0aff)까지의 입구점에 대응하는 경로표를 창조한다.(그림 2)

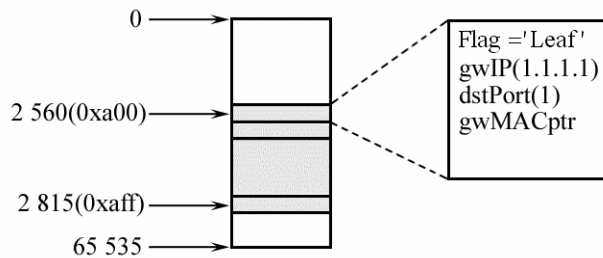


그림 2. 첫 준위표에서 0x0a00부터 0x0aff까지의 입구점에 대응하는 경로표

입구번호 2 560(0x0a00)부터 2 815(0x0aff)(총 256개 입구)까지는 같은 마당값(flag, gwIP, gwMACptr, dstPort, nxtLvlTblPtr...)으로 채워진다.

매 마당값들의 의미는 다음과 같다.

flag: 잎(Leaf)인가 마디(nonLeaf)인가를 나타낸다.

gwIP: 관문IP주소

gwMACptr: 관문MAC주소

dstPort: 목적지포구

nxtLvlTblPtr: 다음준위표에로의 지적자

만일 목적지주소 10.1.1.1을 가진 파킷이 들어올 때 그것의 첫 16bit값은 0x0a01(2 561)이다. 따라서 첫 준위표에서 번호 2 561의 입구가 검사된다. 만일 그 입구가 Leaf마디라면 검색은 끝나며 파킷은 원천지MAC주소와 목적지MAC주소를 갱신한 다음 dstPtr에 대응하는 포구로 경로화된다. 그리고 목적지주소가 10.x.x.x인 모든 파킷들은 그림 2와 같은 경로표에 기초하여 gwIP 1.1.1.1을 가지고 포구 1로 경로화된다.

② 다음 새로운 경로 10.1.1.0/24를 추가한다.

새 경로는 flag마당이 nonLeaf이고 다음준위표에로의 지적자(nxtLvlTblPtr)를 가지고 첫 준위표의 번호 2 561(0x0a01)에 덧쓰기된다.

둘째 준위의 길이가 4bit이므로 다음준위표용으로 16개 입구점을 가진 기억기가 할당된다. 그리고 이 새로운 표의 기초주소는 첫 준위표의 2 561번입구의 nxtLvlTblPtr지적자값으로 설정된다.

새로운 경로의 다음 4bit가 0이므로(0x0a010100의 16—19bit) 둘째 준위표의 첫번째 입구는 또 다른 nonLeaf입구로 리용된다. 둘째 준위표의 모든 입구들은 그것의 부모경로(10.0.0.0/8)의 값과 같은 값으로 채워져야 한다.

망마스크가 24bit로서 첫 준위표와 둘째 준위표의 크기합(16+4)보다 크므로 16개의 입구를 가지는 셋째 준위표가 또 할당되어야 한다. 그리고 새 경로의 다음 4bit(0x0a010100의 20—23bit)가 1이므로 두번째 입구는 새로운 경로용으로 쓰인다. 또한 망마스크가 24bit로서 첫 준위표부터 셋째 준위표까지의 크기합(16+4+4)보다 크지 않으므로 이 입구는 새로운 경로의 Leaf입구로 된다. 그리고 해당한 마당값(gwIP, dstPort 등)들이 그 입구에 채워진다.

목적지IP가 10.1.1.100인 파킷이 들어오면 최종결과를 얻기 위해 다음과 같은 3번의 검색을 거치게 된다.

우선 목적지IP주소의 첫 16bit를 검색하며 이때 결과값은 0x0a01(Non-Leaf)로서 이것은 다음준위검색을 계속하라는것을 의미한다.

다음 목적지IP주소의 다음 4bit를 검색하며 이때 결과값은 0으로서 역시 Non-Leaf로 된다.

끝으로 목적지IP주소의 또 다음 4bit를 검색하며 이때 결과값은 1로 되어 leaf마디를 찾게 되며 검색은 끝난다.

목적지IP주소가 10.1.192.10인 패킷이 들어오면 둘째 준위표에서 leaf마디를 찾게 되며 망주소 10.0.0.0/8의 경로를 얻는다.

목적지IP주소가 10.1.10.10인 패킷의 경우는 셋째 준위표에서 그것의 leaf마디를 찾으며 기대하였던 망주소 10.0.0.0/8의 경로를 얻는다.

이 방법은 Trie경로표보다 표의 추가 및 삭제가 빠른 우점이 있다.

### 3. 실험 및 결과분석

논문에서는 QorIQ계렬처리기 T2080에서 DPAA를 리용한 경로기를 구성하고 LPM경로표를 사용하는 경로기프로그램을 설계하고 매물형리눅스체제에서 실험하였다. 패킷흐름의 처리는 조작체제핵심부내에서가 아니라 응용층에서 직접 DPAA를 리용한 조작으로 실현하여 단위시간당 패킷처리능력을 높이였다. 그리고 2진나무에 기초한 경로표와 Radix-Trie경로표에 기초한 경로탐색시간을 LPM경로표에 의한 경로탐색시간과 비교하였다.

경로표의 종류에 따르는 경로검색시간은 그림 3과 같다. 이로부터 경로표의 크기가 늘어날수록 2진나무나 Radix-Trie나무에 기초한 경로표들에서는 검색시간이 지수함수적으로 늘어나지만 LPM에 기초한 경로표에서는 검색시간이 보다 짧으며 표의 크기가 늘어나도 검색시간에는 영향을 거의 주지 않는다는것을 알수 있다.

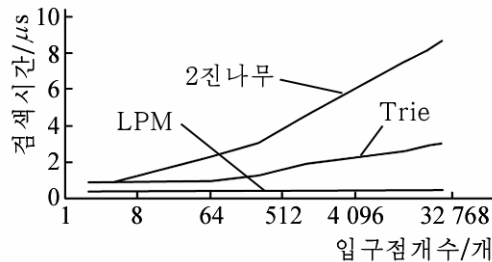


그림 3. 경로표의 종류에 따르는 경로검색시간

그림 4에 DPAA를 리용한 경우와 소켓통신을 리용한 경우의 패킷처리능력을 보여주었다. 그림 4에서 보는바와 같이 소켓통신을 리용한 경우에 최대패킷처리능력은 10kpps를 넘지 못하지만 DPAA를 리용하면 패킷처리능력은 송수신속도에 비례하며 기가비트망경로기의 처리성능을 만족시킬수 있다는것을 알수 있다.

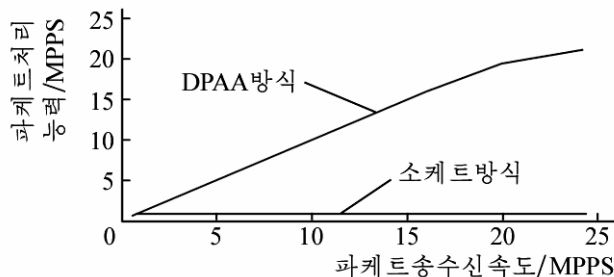


그림 4. DPAA를 리용한 경우와 소켓통신을 리용한 경우의 패킷처리능력

## 참 고 문 헌

- [1] Retvari et al.; IEEE/ACM Transactions on Networking, 24, 149, 2016.
- [2] Ori Rottenstreich; IEEE/ACM Transactions on Networking, 25, 864, 2017.

주체110(2021)년 5월 5일 원고접수

### **Configure Method of the High Speed Router Using DPAA of QorIQ Processor and LPM Routing Table**

*Kim Su Il, Kwak Jin Hyok*

We suggested the method to improve the performance of the router which used the DPAA of the QorIQ T2080 processor, using LPM routing table, and verified its effectiveness by comparing to the method using binary tree or Trie routing table.

Keywords: DPAA, QorIQ processor